

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 20:06:02 ; Search time 18 Seconds
(without alignments)
856.931 Million cell updates/sec

Title: US-10-084-700-2

Perfect score: 1772

Sequence: 1 MTGSNEFKLNQPPEDGIVSV.....DGIFIRQVTDATKPKSPCT 328

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1772	100.0	1 BUB3 HUMAN	O43684 homo sapien
2	1743	98.4	2 BUB3 MOUSE	Q9wv33 mus musculus
3	538	30.4	1 YET7 YEAST	P40066 saccharomyc
4	511.5	28.9	1 RAEL1 SCHPO	P41838 schizosacch
5	509	28.7	1 RAIL1 HUMAN	P78406 homo sapien
6	418.5	23.6	1 YSAK CAEEL	Q93454 caenorhabdi
7	356	20.1	1 BUB3 YEAST	P26449 saccharomyc
8	352.5	19.9	1 RAIL1 ARATH	Q89942 arabidopsis
9	228	12.9	1 YV46 ANASP	Q8y711 anabaena sp
10	209	11.8	1 YS00 ANASP	Q8y7c2 anabaena sp
11	197.5	11.1	1 PKWA THECU	P49495 thermomonos
12	177.5	10.0	1 GBLP CAEEL	Q20836 caenorhabdi
13	176	9.9	1 HET1 PODAN	Q00808 podospora a
14	175.5	9.9	1 PRL1 ARATH	Q42384 arabidopsis
15	173	9.8	1 YH92 CAEEL	Q23256 caenorhabdi
16	171.5	9.7	1 GBL1 RAT	P54311 rattus norv
17	170.5	9.6	1 WDR1 CHICK	Q93277 gallus gall
18	169.5	9.6	1 GBL1 HUMAN	P04901 homo sapien
19	169.5	9.6	1 T2D2 HUMAN	Q15542 homo sapien
20	169.5	9.6	1 PWP2 SCHPO	Q9c1x1 schizosacch
21	169	9.5	1 PRL2 ARATH	Q91910 arabidopsis
22	169	9.5	1 WDR1 XENLA	Q9w7f2 xenopus lae
23	168.5	9.5	1 GBLP ARATH	Q24456 arabidopsis
24	167	9.4	1 GBLP TRYBB	Q94775 trypanosoma
25	167	9.4	1 PRL1 YEAST	P40968 saccharomyc
26	166	9.4	1 GBLP CHLRE	P25387 chlamydomon
27	166	9.4	1 RCO1 NEUCR	P78706 nematospora
28	165.5	9.3	1 YJL2 YEAST	P47025 saccharomyc
29	164.5	9.3	1 GBLP LEICH	P27434 leishmania
30	164.5	9.3	1 WU11 SCHPO	Q09715 schizosacch
31	164	9.3	1 T146 HUMAN	Q9c018 homo sapien
32	163.5	9.2	1 GBLP LEIMA	Q25306 leishmania
33	163.5	9.2	1 YL24 ANASP	Q9yvw57 anabaena sp

RESULT 1

ID	BUB3 HUMAN	STANDARD;	PRT;	328 AA.
AC	O43684; O43685;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Mitotic checkpoint protein BUB3.			
GN	BUB3.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Spleen, and Testis;			
RA	Seeley T.W.;			
RL	Submitted (PDB-1998) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98327111; PubMed=9660858;			
RA	Taylor S.S., Ha E., McKeon F.;			
RT	"The human homologue of Bub3 is required for kinetochore localization of Bub1 and a Mad3/Bub1-related protein kinase.";			
RL	J. Cell Biol. 142:1-11(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Chan G.K.T., Yen T.J.;			
RT	"Human Bub3 is a kinetochore protein that interacts with hBub1 and hBUBR1.";			
RL	Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Spleen, and Testis;			
RA	Seeley T.W.;			
RT	"hUBUB3, ligand for hUBUB1 mitotic checkpoint kinase.";			
RL	Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain, and Placenta;			
RX	MEDLINE=2238257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Bhat N.K., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshitoki S., Carninci P., Prange C., Raha S.S., McQuillan N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., LeMwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			

P93340 nicotiana p
P49026 nicotiana t
P11017 bos taurus
P79959 xenopus lae
P11016 homo sapien
P90648 dictyosteli
Q93134 biophalari
Q21215 caenorhabdi
Q24076 medicago sa
Q39836 glycine max
P17343 caenorhabdi

ALIGNMENTS

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- FUNCTION: REQUIRED FOR KINETOCHORE LOCALIZATION OF BUB1.
 CC -!- SUBUNIT: INTERACTS WITH BUB1 AND BUBR1.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Contains 5 WD repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF047472; AAC28438.1; -
 CC EMBL; AF053304; AAC06258.1; -
 CC EMBL; AF081496; AAC36307.1; -
 CC EMBL; AF047473; AAC28439.1; -
 CC EMBL; BC005138; AAH05138.1; -
 CC EMBL; BC022438; AAH22438.1; -
 CC GenBank; HGNC:1151; BUB3.
 CC MIM; 603719; -
 CC GO; GO:0005699; C:kinetochore; TAS.
 CC GO; GO:0008283; P:cell proliferation; TAS.
 CC GO; GO:0007067; P:mitosis; TAS.
 CC InterPro; IPR001680; WD40.
 CC Pfam; PF00400; WD40; 5.
 CC PRINTS; PR00320; GPROTEINBRPT.
 CC ProDom; PD000018; WD40; 1.
 CC SMART; SM00320; WD40; 6.
 CC PROSITE; PS00678; WD_REPEATS_1; FALSE_NEG.
 CC PROSITE; PS50082; WD_REPEATS_2; 2.
 CC PROSITE; PS50294; WD_REPEATS_REGION; 1.
 CC Nuclear protein; Repeat; WD repeat.
 CC REPEAT 5 43
 CC FT REPEAT 46 83
 CC FT REPEAT 86 124
 CC FT REPEAT 128 163
 CC FT REPEAT 223 262
 CC FT CONFLICT 156 165
 CC FT CONFLICT 165 165
 CC FT CONFLICT 270 274
 CC MISSING (IN REF. 4).
 CC SEQUENCE 328 AA; 37155 MW; 2915572A57368B5A CRC64;
 CC -----
 CC Query Match 100.0%; Score 1772; DB 1; Length 328;
 CC Best Local Similarity 100.0%; Pred. No. 5.9e-146;
 CC Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 1 MTGSEFKLNQPPEDGSSVVFSPNTSQFLVSSWDTSVRLYDVPANSMRLKYQHTGAVL 60
 CC DB 1 MTGSEFKLNQPPEDGSSVVFSPNTSQFLVSSWDTSVRLYDVPANSMRLKYQHTGAVL 60
 CC QY 61 DCAFDPTPHAWSGGLDQHLKMDLNTDQENLVGTHDAPIRVCVEYCPVNVVMTGSDQTV 120
 CC DB 61 DCAFDPTPHAWSGGLDQHLKMDLNTDQENLVGTHDAPIRVCVEYCPVNVVMTGSDQTV 120
 CC QY 121 KLWDRTPCNACTFSQPEKVTYLSVSGDRLVGTAGRRVLVWDLNNKGYVQQRRESISKY 180
 CC DB 121 KLWDRTPCNACTFSQPEKVTYLSVSGDRLVGTAGRRVLVWDLNNKGYVQQRRESISKY 180
 CC QY 181 QTRCTRAPKNGCYVLSSTEGRAVVEYLDPSPVQKKVAFKCHRLKNNIEQIYPPVNAI 240
 CC DB 181 QTRCTRAPKNGCYVLSSTEGRAVVEYLDPSPVQKKVAFKCHRLKNNIEQIYPPVNAI 240
 CC QY 241 SFNNIHTFATGSGDFVNIWDPFNKRLCQPHRYPTSIASLAFSNDGTTLAIASSYMYE 300
 CC DB 241 SFNNIHTFATGSGDFVNIWDPFNKRLCQPHRYPTSIASLAFSNDGTTLAIASSYMYE 300
 CC QY 301 MDDTEHPDGIFIRQVTDATETPKSPCT 328
 CC DB 301 MDDTEHPDGIFIRQVTDATETPKSPCT 328

RESULT 2
 BUB3 MOUSE
 ID BUB3 MOUSE STANDARD; PRT; 326 AA.
 AC Q9WV43; P97397;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Mitotic checkpoint protein BUB3 (WD-repeat type I transmembrane
 DE protein A72.5).
 GN BUB3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=99342046; PubMed=10411903;
 RA Downes A., Xie X., Yan W., Li J., Palacios R.;
 RA Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: REQUIRED FOR KINETOCHORE LOCALIZATION OF BUB1.
 CC -!- SUBUNIT: INTERACTS WITH BUB1 AND BUBR1.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- SIMILARITY: Contains 5 WD repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF149822; AAD38038.1; -
 CC EMBL; U67327; AAB39606.1; -
 CC MGD; MGI:1343463; BUB3.
 CC InterPro; IPR001680; WD40.
 CC Pfam; PF00400; WD40; 5.
 CC PRINTS; PR00320; GPROTEINBRPT.
 CC ProDom; PD000018; WD40; 1.
 CC SMART; SM00320; WD40; 6.
 CC PROSITE; PS00678; WD_REPEATS_1; FALSE_NEG.
 CC PROSITE; PS50082; WD_REPEATS_2; 2.
 CC PROSITE; PS50294; WD_REPEATS_REGION; 1.
 CC Nuclear protein; Repeat; WD repeat.
 CC REPEAT 5 43
 CC FT REPEAT 46 83
 CC FT REPEAT 86 124
 CC FT REPEAT 128 163
 CC FT REPEAT 223 262
 CC FT CONFLICT 156 156
 CC FT CONFLICT 165 165
 CC FT CONFLICT 270 274
 CC CQFHR -> LFVPS (IN REF. 2).
 CC SEQUENCE 326 AA; 36984 MW; F5B80510B82A1342 CRC64;
 CC -----
 CC Query Match 98.4%; Score 1743; DB 1; Length 326;
 CC Best Local Similarity 99.7%; Pred. No. 1.9e-143;
 CC Matches 324; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC QY 1 MTGSEFKLNQPPEDGSSVVFSPNTSQFLVSSWDTSVRLYDVPANSMRLKYQHTGAVL 60
 CC DB 1 MTGSEFKLNQPPEDGSSVVFSPNTSQFLVSSWDTSVRLYDVPANSMRLKYQHTGAVL 60
 CC QY 61 DCAFDPTPHAWSGGLDQHLKMDLNTDQENLVGTHDAPIRVCVEYCPVNVVMTGSDQTV 120

Db 61 DCAFYDPTAHSGGLDHLQKMDLNTDQENLVGTHDAPIRCVVEYCPVNVVWVTSNDQTV 120
 QY 121 KLMDPRTPCNAGTFSQPEKVTYLSVSGDLIVGTAGRRVLVWDLNMGVYQORRESSLKY 180
 Db 121 KLMDPRTPCNAGTFSQPEKVTYLSVSGDLIVGTAGRRVLVWDLNMGVYQORRESSLKY 180
 QY 181 QTRCIRAFNKGQYVLSSTEGRAVVEYLDPSPEVOKKKYAPKCHKLKNNIEQIYVNAI 240
 Db 181 QTRCIRAFNKGQYVLSSTEGRAVVEYLDPSPEVOKKKYAPKCHKLKNNIEQIYVNAI 240
 QY 241 SFNHNHTATGSGDGFVNWDPNFKKRLCQPHRYPTSIASLAFSNDGTTLAIASSYME 300
 Db 241 SFNHNHTATGSGDGFVNWDPNFKKRLCQPHRYPTSIASLAFSNDGTTLAIASSYME 300
 QY 301 MDDTEHPDGIFIRQVTDATPKS 325
 Db 301 MDDTEHPDGIFIRQVTDATPKS 325

RESULT 3

YEST_YEAST STANDARD; PRT; 365 AA.
 AC P40066;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hypothetical 40.5 kDa Tri-Asp repeats containing protein in
 DE NUP157-PH5 intergenic region.
 GN YER107C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Schizosaccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX PubMed=9169868;
 RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
 RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
 RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
 RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
 RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
 RL Nature 387:78-81 (1997).
 CC -1- SIMILARITY: BELONGS TO THE RAE1 FAMILY OF WD-REPEAT PROTEINS.
 CC -1- SIMILARITY: CONTAINS 4 WD REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>;
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U18839; AAB64662.1; -;
 DR PIR; S50610;
 DR SGD; S000909; YER107C.
 DR GO; GO:0005643; C:nuclear pore; IDA.
 DR GO; GO:0006406; P:nuclear membrane organization and biogenesis; IMP.
 DR GO; GO:0006998; P:nuclear pore organization and biogenesis; IMP.
 DR GO; GO:0006999; P:nuclear pore organization and biogenesis; IMP.
 DR GO; GO:0006999; P:nuclear pore organization and biogenesis; IMP.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 4.
 DR PRINTS; PR03020; GPROTEINBRPT.
 DR SMART; SMC0320; WD40; 4.
 DR PROSITE; PS00678; WD_REPEATS_1; FALSE_NEG.
 DR PROSITE; PS00082; WD_REPEATS_2; 4.
 DR PROSITE; PS00294; WD_REPEATS_REGION; 1.
 KW Hypothetical protein; Repeat; WD repeat.

FT REPEAT 34 74 WD 1.
 FT REPEAT 78 117 WD 2.
 FT REPEAT 119 162 WD 3.
 FT REPEAT 267 306 WD 4.
 SQ SEQUENCE 365 AA; 40522 MW; 3F5B201EB1BAEC8 CRC64;
 Query Match 30.4%; Score 538; DB 1; Length 365;
 Best Local Similarity 35.7%; Pred. No. 3.7e-39;
 Matches 122; Conservative 59; Mismatches 135; Indels 26; Gaps 10;
 QY 4 SNFKLNQPPEDGIGSSKPFSPNTSOFLLVSSWDTSVRLVYDVNPANSLRYKH-TGAVLDC 62
 Db 26 ANDIVNSPAEDSISDIAPFSQDQDFMSASDCKVRIWDVQVQFQCAQHESSPVL 85
 QY 63 AFY--DPTHAWSGGLDHLQKMDLNTDQENLVGTHDAPIRCVVEY--CPVNV--WVTGSM 116
 Db 86 TRWSNDGTVKASGGCDNALKLYDIAGGQTCQIQGMHSAPIKVLRFVOCGGSNTECIVTGSN 145
 QY 117 DOTVGLMDPRTPCNAGTFSQPEKVTYLSVSGDLIVGTAGRRVLVWDLNMGVYQORRES 176
 Db 146 DTKIKYWDMEQPOPVSTVMMPEVYSGMDNKQSLVAVATAERHAIINLANPTTFKATTS 205
 QY 177 SLKYOTRCIRAFNKGQYVLSSTEGRAVVEYLDPSPEVOKKK-YAPKCHRLK-----E 228
 Db 206 PLKWQTRCVACVNEADGAYIGSVGRCIRYIDG--MOKSGFSFCHRONPNRAPS 263
 QY 229 NNIEQIYVNAISFNHNHTATGSGDGFVNWDPNFKKRLCQPHRYPTSIASL---AFS 285
 Db 264 NGQSLVYVNSIAFHPLVGTFTVAGDGTFFNPKXQNRHL---KGYPQLQASIPVCSFN 320
 QY 286 NDTGTLAIASSYMY---EMDTEHPDGIFIRQVTDATPK 324
 Db 321 RNSGVFAVALSYDMHQHGMGNREDYFNVRILHATTDVEVKEK 362

RESULT 4

RAE1_SCHPO STANDARD; PRT; 352 AA.
 AC P41838;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Poly(A)+ RNA export protein.
 DE RAE1 OR SPBC16A3.05C.
 GN Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95221400; PubMed=7706287;
 RA Brown J.A., Bharathi A., Ghosh A., Whalen W., Fitzgerald E.,
 RA Dhar R.;
 RT "A mutation in the Schizosaccharomyces pombe rae1 gene causes defects
 RT in poly(A)+ RNA export and in the cytoskeleton.";
 RL J. Biol. Chem. 270:7411-7419 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidaigo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,


```

RESULT 6
ID YSAK_CABEL STANDARD; PRT; 373 AA.
YSAK_CABEL
ID YSAK_CABEL PRT; 373 AA.
AC Q93454;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 41.4 kDa Trip-Asp repeats containing protein F10G8.3 in
DE chromosome 1.
GN F10G8.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
ON NCBI_TaxID=6239;
RX [1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Basham V.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 4 WD repeats.
CC -1- SIMILARITY: BELONGS TO THE RAE1 FAMILY OF WD-REPEAT PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z80216; CAB02280.1; -.
DR F10G8.3;
DR F10G8.3;
DR WormPep; F10G8.3; CE093338.
DR InterPro; IPRO01680; WD40.
DR Pfam; PF00400; WD40; 4.
DR PRINTS; PR00320; GPROTEINRPT.
DR SMART; SM00320; WD40; 4.
DR PROSITE; PS00678; WD_REPEATS 1; 2.
DR PROSITE; PS00082; WD_REPEATS 2; 1.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Repeat; WD repeat.
FT REPEAT 40..82
FT REPEAT 87..126
FT REPEAT 128..169
FT REPEAT 276..315
FT REPEAT 315..WD 4.
SQ SEQUENCE 373 AA; 41413 MW; B0B151DB47FC3526 CRC64;
Query Match 23.6%; Score 418.5; DB 1; Length 373;
Best Local Similarity 31.5%; Pred. No. 8.3e-29;
Matches 109; Conservative 63; Mismatches 145; Indels 29; Gaps 11;
QY 2 TGSNEFLKNQPEDGIGSSVKRSPNTSQ--FLVSSWDTSVRLYDV-PANSMRLKYQHT-- 56
Ddb 30 TQNDFLVDGAPEDIQVIKFSPTQDPKMLACGSDWGTIRVMMFNDAFTFGKAQNTIP 89
QY 57 GAVLDCAFY-DPTHAWSGGLDHLKMDLNTDQNLVGTHTDAPIRCVCEYCPVN----V 110
Ddb 90 APILDIAMIEDSSKFIACADKEARLWDLASQNAVVGTHDGFVKT--CHWINGNNYQC 146
QY 111 MVTSGWDQTVKLNDRPTCNAGTSQ----PEKVYTLVSGDRLIVGTAGRRVLVMDLRNM 167
Ddb 147 LMTGSFDTLRFDMKNLPNQTQMAQIQLPERVYAADVLPMAVVALANKHKIVYNLENG 206
QY 168 GYVQQRRESSLKYOTRCIRAFPNKQ----CYVLSSIEGRVAVEVLDPSPVQKKKYAPK 222
Ddb 207 PTEVKDIESQKLFQIRCLISIFKSNQNPAGFALGSIAGRVAQVQVDVAN--PKDNFTFK 264
QY 223 CHRLKE--NNIEQIYPVNAISFNHINHTFATGSGDGFVNIWDPFNKRLCQFHYETSLA 280
Ddb 265 CHRSaelVNGQEIVAVNDICFHPQHGTTLVIGSDGRYSMDKDKARTYKLTSEHPHPLT 324
QY 281 SLAFSNDGTTLAIASSYWNEMOOTEHPEDG---IFIRQVTDATPK 323

```


OS Anabaena sp. (strain PCC 7120).
 CC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2159285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium Anabaena sp. strain PCC 7120".
 RL DNA Res. 8:205-213(2001).
 CC -!- SIMILARITY: Contains 1 NB-ARC domain.
 CC -!- SIMILARITY: Contains 15 WD repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AP003590; BAB74499.1; -;
 DR PIR; A12155; A12155.
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR002182; NB-ARC.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00931; NB-ARC; 1.
 DR Pfam; PF00400; WD40; 14.
 DR PRINTS; PR00364; DISEASEREST.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR ProDom; PD000018; WD40; 1.
 DR SMART; SM00320; WD40; 14.
 DR PROSITE; PS00678; WD REPEATS 1; 9.
 DR PROSITE; PS00082; WD REPEATS 2; 14.
 DR PROSITE; PS00294; WD REPEATS REGION; 1.
 KW Hypothetical protein; Repeat; WD repeat; Complete proteome.
 FT REPEAT 55 93
 FT REPEAT 640 679
 FT REPEAT 682 721
 FT REPEAT 724 763
 FT REPEAT 766 807
 FT REPEAT 809 849
 FT REPEAT 850 889
 FT REPEAT 892 931
 FT REPEAT 934 975
 FT REPEAT 976 1017
 FT REPEAT 1019 1059
 FT REPEAT 1060 1101
 FT REPEAT 1103 1143
 FT REPEAT 1144 1183
 FT REPEAT 1186 1227
 SQ SEQUENCE 1258 AA; 139513 MW; 45DF03B91170C451 CRC64;
 Query Match 11.8%; Score 209; DB 1; Length 1258;
 Best Local Similarity 24.5%; Pred. No. 5.4e-10;
 Matches 78; Conservative 60; Mismatches 139; Indels 42; Gaps 11;
 QY 15 DGISVSKSPNTSQFLVSSMDTSVRLYDV-PANSMRLKYQHTGAVLDCAP-YDPTHAWS 72
 Db 769 DVMRCVAFSPD-GNTLASSAADHTIKLMDVSGKCLRTLKSHTCWVRVAFSADGQTLAS 827
 QY 73 GGLDHLQKMDLNTQ--ENLVGTHDAPICVCEPVNVNVTGSGWQTVKLMDFRTP-C 129
 Db 828 GSGDRTIKIWNVHTGECLTKTYIG-HTNSVYSIATSPSKILVSGSGDRITKLMDCQTHIC 886
 QY 130 NAGTSQPKKVVTLVS--GRLIVGTAGRVLYVNLN-----MGYVQ 171
 Db 887 IKTLGHTNEVCSVAFSPDQQTACVSLDQSVRLNCRCTGQCLKAWYNTDMLPVAFSP 946

QY 172 QRRESSLKQTRCIRAPNKGQVLSIEGRVAVEY-LDPSPEVQKKKAYF----- 221
 Db 947 DRQILASGSDKTKVLDWQTKYIISLEGTFTFYGIAFSPDSQTLASSTSSVRLWN 1006
 QY 222 ----KCHRLKENNIEQIYPVNAISPHNIHNTFATGSDGFVNIDPFNKKRLCQPHRYPT 277
 Db 1007 ISTQCQFILLEHTDWY---AVVFHPQGIATGSDACTVKLWNISTGQCLTKLSEHSD 1063
 QY 278 SIASLAFSNDGTTLAIASS 296
 Db 1064 KILGWAMSPDQQLLASASA 1082

RESULT 11
 PKWA_THECU STANDARD; PRT; 742 AA.
 ID P49695;
 AC DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Probable serine/threonine-protein kinase pKWA (EC 2.7.1.37).
 GN PKWA OR PKW1.
 OS Thermomonospora curvata.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptosporangineae; Thermomonosporaceae; Thermomonospora.
 OX NCBI_TaxID=2020;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CCM 3352;
 RX MEDLINE=96200125; PubMed=8631732;
 RA Janda L., Tichy P., Spizek J., Petricek M.;
 RT "A deduced Thermomonospora curvata protein containing
 RT serine/threonine protein kinase and WD-repeat domains.";
 RL J. Bacteriol. 178:1487-1489(1996).
 CC -!- FUNCTION: MAY PLAY A REGULATORY ROLE DURING THE COMPLEX GROWTH
 CC CYCLE AND IN SECONDARY METABOLITE PRODUCTION.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -!- SIMILARITY: Contains 7 WD repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF115313; AAB05822.1; -;
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_Thr_kinase.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00400; WD40; 7.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR ProDom; PD000018; WD40; 5.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00678; WD REPEATS 1; 5.
 DR PROSITE; PS00082; WD REPEATS 2; 7.
 DR PROSITE; PS00294; WD REPEATS REGION; 1.
 KW Transferase; Serine/Threonine-protein kinase; ATP-binding; Repeat;
 WD repeat.
 KW DOMAIN 16 266 PROTEIN KINASE.
 FT NP BIND 22 30 ATP (BY SIMILARITY).
 FT BINDING 44 44 ATP (BY SIMILARITY).
 FT ACT_SITE 138 138 BY SIMILARITY.
 FT DOMAIN 301 394 PRO/GRU/ASP-RICH (SPACER).
 FT REPEAT 455 496 WD 1.
 FT REPEAT 497 538 WD 2.

FT REPEAT 539 580 WD 3.
 FT REPEAT 581 621 WD 4.
 FT REPEAT 622 663 WD 5.
 FT REPEAT 664 705 WD 6.
 FT REPEAT 706 742 WD 7.
 FT SEQUENCE 742 AA; AC1734640DB4383D CRC64;

Query Match 11.1%; Score 197.5; DB 1; Length 742;
 Best Local Similarity 27.1%; Pred. No. 2.7e-09;
 Matches 70; Conservative 48; Mismatches 112; Indels 23; Gaps 10;

QY 15 DGSSVFSPTQGLVSSWTSVRLYDVPANSMRLKYQ-HTGAVLDCAPY-DPTHAWS 72
 DB 502 DWVRAVAFSD-GALLASGSDATVRLMDVAABEAVFEGHTHYVLDIAFSPDGMVAS 560
 QY 73 GGLDHQKMLNDTDOENLY-GTHDAPIRCVCEYCPVNNVMTGSDQTVKLDWPTPCNA 131
 DB 561 GSRDGTARLNNVATGTEHAKVKGHTDYVVAFAFSDGMVASGSRDGTIRLMDVATGKER 620
 QY 132 GTFSPQ-EKVTYLSVSGD-RLIVGTAGRRVLVMDLRNMGVQORRESSLKYOTRCIRAPP 189
 DB 621 DVLOAPENVSLAFSPDGSMLVHSDSTVHLMDVASGEALH-----TFEGHTDWRAVA 675
 QY 190 -NKQGYVLSSTIEGRVAVVEYLDPSPEVQKKYAFKCHRLKENNIEQIYPNNAISFNHNT 248
 DB 676 FSPDGLASGSDORTIRLMDVAABE-----HTTLEGHTE---PVHSAFAHPEGTT 724
 QY 249 FATGSGDGFVNIW 261
 DB 725 LASASEDGTIRW 737

Query Match 10.0%; Score 177.5; DB 1; Length 356;
 Best Local Similarity 25.2%; Pred. No. 5.9e-08;
 Matches 70; Conservative 40; Mismatches 143; Indels 25; Gaps 8;

QY 30 LLVSSMDTSVRLYD-VPANSMRLKYQHTGAVLDCAPYDPTHAWS-GGLDHQ-----LKM 82
 DB 83 IVSSSDGKVIWMDGFTTNKEHALTMTPTTWMACAFSPSSQMIACGLDNKCSWPLSFE 142
 QY 83 DLNTDQENLVGTHDAPIRCVCEYCPVNNVMTGSDQTVKLDWPTPCNAGTF-SQEKVY 141
 DB 143 DDIIQKKRQVATHTSYNSCCFTLRSDNLLITGSDGTCAIMDVESGQLIQNHFGHTGDFV 202
 QY 142 TLSV-----SGDLVGTAGRRVLVMDLRNMGVQORRESSLKYOTRCIRAPPNKGVYLS 197
 DB 203 AIDVPKCDTGNFTISAGADKHSVLWDIRSGCQVQSPGEGHEADINT--VRPHNGDAFATG 260
 QY 198 SIEGRVAVVEYLDPSPEVQKKYAFKCHRLKENNIEQIYPNNAISFNHNTFTGSDGDF 257
 DB 261 SDDATCRLFLDLRADQV-----CVYKES---ILFPVNGVDFSLSGRILFAGYGDYR 309
 QY 258 VNIWDPFNKRLCQFHRYPTYSIASLAFSNDGTTLAIAS 295
 DB 310 VGWDSLKCAHSHVLYCHENRISCLRTSPDGTAVCSAS 347

RESULT 13
 HET1_PODAN
 ID_HET1_PODAN STANDARD; PRT; 1356 AA.
 AC_Q00808;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Vegetatable incompatibility protein HET-B-1.
 GN HET-B-1.
 OS Podospora anserina.
 OC Eukaryota, Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Lasiosphaeriaceae; Podospora.
 NCBI_TaxID=5145;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=96009991; PubMed=7557402;
 RA Saupé S., Turcq B., Begueret J.
 RT "A gene responsible for vegetative incompatibility in the fungus Podospora anserina encodes a protein with a GTP-binding motif and G beta homologous domain."
 RT modified and this statement is not removed. Usage by, and for commercial

RESULT 12
 GBB2_CAEEL
 ID_GBB2_CAEEL STANDARD; PRT; 356 AA.
 AC_Q00636;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Guanine nucleotide-binding protein beta subunit 2.
 GN GBB-2 OR F52A8.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Periderinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN [1]
 RN SEQUENCE FROM N.A., AND INTERACTIONS.
 RP STRAIN=Bristol N2;
 RX MEDLINE=21231612; PubMed=11333232;
 RA van Der Linden A.M., Simmer F., Cuppen E., Plasterk R.H.A.;
 RT "The G protein beta subunit gbb-2 in Caenorhabditis elegans regulates the (G)alpha-G(q)alpha signaling network through interactions with the regulator of G protein signaling proteins egl-10 and eat-16.";
 RL Genetics 158:221-235(2001).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Bristol N2;
 RX Wilkinson J.;
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE INVOLVED AS A MODULATOR OR TRANSDUCER IN VARIOUS TRANSMEMBRANE SIGNALING SYSTEMS. THE BETA AND GAMMA CHAINS ARE REQUIRED FOR THE GTPASE ACTIVITY, FOR REPLACEMENT OF GDP BY GTP, AND FOR G PROTEIN-EFFECTOR INTERACTION.
 CC -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA AND GAMMA). Interacts with goa-1, eat-16, egl-10, egl-30.
 CC -!- SIMILARITY: Contains 7 WD repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by, and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AP291847; AAK55964.1;
 CC EMBL; Z71263; CAA95824.1;
 CC PIR; T22478; T22478.
 CC DR HSSP; P04901; 1TBG.
 CC DR W01Pep; F52A8-2; CE17845.
 CC DR InterPro; IPR001632; Gprotein_B.
 CC DR InterPro; IPR001680; WD40.
 CC Pfam; PF00400; WD40; 7.
 CC DR PRINTS; PR00319; GPROTEINB.
 CC DR PRINTS; PR00320; GPROTEINBPT.
 CC DR ProDom; PD000018; WD40; 3.
 CC DR SMART; SM00320; WD40; 7.
 CC DR PROSITE; PS00678; WD_REPEATS_1; 3.
 CC DR PROSITE; PS00882; WD_REPEATS_2; 5.
 CC DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
 CC KW Transducer; Repeat; WD repeat; Multigene family.
 FT REPEAT 66 96 WD 1.
 FT REPEAT 108 138 WD 2.
 FT REPEAT 154 184 WD 3.
 FT REPEAT 196 228 WD 4.
 FT REPEAT 240 270 WD 5.
 FT REPEAT 284 314 WD 6.
 FT REPEAT 326 356 WD 7.
 FT SEQUENCE 356 AA; 39453 MM; 293AA6DCF317D832 CRC64;

```
RL Gene 162135-139(1995).
CC -1- FUNCTION: RESPONSIBLE FOR VEGETATIVE INCOMPATIBILITY THROUGH
CC SPECIFIC INTERACTIONS WITH DIFFERENT ALLELES OF THE UNLINKED GENE,
CC HET-C.
CC -1- SIMILARITY: Contains 10 WD repeats.
CC -1- SIMILARITY: Contains 1 NACHT domain.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; L28125; AAA85775.1; -.
DR PIR; T18521.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR001680; WD40.
DR Pfam; PF004400; WD40; 10.
DR PRINTS; PR00320; GPROTEINRPT.
DR ProDom; PD000018; WD40; 10.
DR SMART; SM00320; WD40; 10.
DR PROSITE; PS00837; NACHT; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 10.
DR PROSITE; PS00082; WD_REPEATS_2; 10.
DR PROSITE; PS02294; WD_REPEATS_REGION; 1.
KW GTP-binding; Repeat; WD repeat.
FT DOMAIN 294 629
FT NP_BIND 300 307 GTP (POTENTIAL).
FT REPEAT 839 869 WD 1.
FT REPEAT 881 911 WD 2.
FT REPEAT 923 953 WD 3.
FT REPEAT 965 995 WD 4.
FT REPEAT 1007 1037 WD 5.
FT REPEAT 1049 1079 WD 6.
FT REPEAT 1091 1121 WD 7.
FT REPEAT 1133 1163 WD 8.
FT REPEAT 1175 1205 WD 9.
FT REPEAT 1217 1247 WD 10.
SQ SEQUENCE 1356 AA; 149765 MW; 965FB319844E0651 CRC64;

Query Match
Best Local Similarity 22.9%; Score 176; DB 1; Length 1356;
Matches 69; Conservative 55; Mismatches 119; Indels 58; Gaps 12;

QY 19 SVKPSNTSQFLVSSWDTSLVRLYDVPANMRKYQ-HTGAVLDCAFY-DPTHANSGGLD 76
DB 888 SVAFSPDRER-VASGSGDDTKIKINDAASGCTQTLEGGGRVQSVAFSPDQGVASGSD 946
QY 77 HOLKMHDL--NTDQENLVGTHDAPICRVCPEVNVMTGSDQTVKLDPRTPCNAGTF 134
DB 947 HTIKINDAASGCTQTLEGGHSSVLSVAFSPDQGVASGSDTKIKINDTA-----SGTC 1001
QY 135 SQP-----EKVYTLVS--GDRLLVGTAGRLVLDNRNMVGVQORRESSLYQTRCIRA 187
DB 1002 QTLEGGHSSVLSVAFSPDQGVASGSDTKIKINDTA-----GTCQT 1046
QY 188 FPKQGYVLSSEIGRAVVEYLDPSPEVQKKYAFKCHLK-----ENNIOI 234
DB 1047 LEHGGGWQSVVF-----SPDQGVASGSDDTIKINDAVSGTCTQTLEGGHSDV 1096
QY 235 YPNATSEPHNHTATGSGDGFVNITDPFNKRLCPHRYFTSLAFSNDGTTLAIA 294
DB 1097 W----SVAFSPDQGVASGSDGTIKINDAASGCTQTLEGGHGWVHVSFAFSPDQGVASG 1153
QY 295 S 295
DB 1154 S 1154
```

RESULT 14

PRLI_ARATH

```
ID AC Q42384;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE PPI/PP2A phosphatases pleiotropic regulator PRLI.
OS PRLI OR AT4G15900 OR DL3990W.
GN Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Nemeth K., Putnok P., Stankovic B., Bako L., Mathur J., Redei G.P.,
RA Schell J., Koncz C.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=98121113; PubMed=9461215;
RA Bevan M., Bancroft I., Bent E., Love K., Goodman H.M., Dean C.,
RA Bergkamp R., Dirkse W., van Staveren M., Stiekema W., Drost L.,
RA Ridley P., Hudson S.-A., Patel K., Murphy G., Piffanelli P.,
RA Wedler H., Wedler E., Wambutt R., Weitzenegger T., Pohl T., Terryn N.,
RA Gielens J., Villarroel R., De Clercq R., van Montagu M., Lecharny A.,
RA Aubourg S., Gy I., Kreis M., Lao N., Kavanagh T., Hempel S.,
RA Kottler P., Entian K.-D., Rieger M., Schaefer M., Funk B.,
RA Mueller-Auer S., Silvey M., James R., Monfort A., Pons A.,
RA Puigdomenech P., Douka A., Vouklatou E., Milloni D., Hatzopoulos P.,
RA Piravandi E., Obermaier B., Hilbert H., Duesterhoft A., Moores T.,
RA Jones J.D.G., Eneva T., Palme K., Benes V., Rechmann S., Ansgore W.,
RA Cooke R., Berger C., Delseny M., Voet M., Volckaert G., Mewes H.-W.,
RA Klosterman S., Schueller C., Chalwatzis N.;
RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of
RL Arabidopsis thaliana.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansgore W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Moeljan P., Klein Lankhorst R., Rose M., Hauf J., Kottler P.,
RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buysnaert C., Gielens J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark I., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA Neumann S., Argirou A., Vitale D., Liguori R., Piravandi E.,
RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Cheddor F., Cooke R., Berger C., Monfort A., Cascuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Bauges M., Terol J., Torres A.,
RA Perez-Perez A., Furnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
```

```
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Speth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Marra M., Marijnsen R., McCombie W.R.,
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
thaliana"; 769-777(1999).
RL Nature 402:769-777(1999).
CC -!- FUNCTION: PLEIOTROPIC REGULATOR OF PP1 AND PP2A PHOSPHATASES.
CC -!- SIMILARITY: Contains 7 WD repeats.
CC -!- SIMILARITY: BELONGS TO THE PRL1/PRL2 FAMILY OF WD-REPEAT PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X82825; CAA58032.1; --
CC EMBL; X82824; CAA58031.1; --
CC EMBL; Z97339; CAB10369.1; --
CC EMBL; AL161542; CAB78632.1; --
CC FIRM; S49820; S49820.
CC InterPro; IPR001680; WD40.
CC PRINTS; PR00320; GPROTEINBRPT.
CC ProDom; PD000018; WD40; 2.
CC SMART; SM00320; WD40; 7.
CC PROSITE; PS00678; WD_REPEATS_1; 2.
CC PROSITE; PS50082; WD_REPEATS_2; 5.
CC PROSITE; PS50294; WD_REPEATS_REGION; 1.
CC Repeat; WD repeat.
CC REPEAT 174 204 WD 1.
CC REPEAT 216 246 WD 2.
CC REPEAT 258 288 WD 3.
CC REPEAT 300 330 WD 4.
CC REPEAT 342 371 WD 5.
CC REPEAT 384 413 WD 6.
CC REPEAT 433 463 WD 7.
CC SEQUENCE 486 AA; 54009 MW; 604D3E6FCD8A998 CRC64;
Query Match 9.9%; Score 175.5; DB 1; Length 486;
Best Local Similarity 21.7%; Pred. No. 1.2e-07;
Matches 71; Conservative 63; Mismatches 124; Indels 69; Gaps 16;
QY 17 ISSVKFSPNTSQFLVSSWDTSVRLYVPANSMRLKYQ-HTGAVLPCAFYD-PTHAWSGG 74
DB 179 VRSVAFDP-:::-----DKTADRTIKWDVATGLVLTGTHIEQVGRGLAVSNRHTYMFSA 237
QY 75 LDHQLKMDLNTDQENLVGT---HDAPIRCVCEYCPVNVAVTGSWDTVKLMDPRPCN- 130
DB 238 DDQVKQKDWL---EQNKVRSYHGHISGVYCLAUPTLDVLTGDSVCRVWMDIRYMQI 295
QY 131 -----AGTSQEKVYVTLVSGDRLIVGTAGRRVLVWDLRNNGYVQQRRESSLKY 180
DB 256 FALSGHONTVCVSTRP-----TDPQVVGTGSHDTTIKFMDLR-----YGKTMSTLTH 342
QY 181 QTRCIRAF---PNQGVVLSIEGRVAVEIDLPSPVQKKYAFK---CHRLKENNIEQ 233
DB 343 HKKSVMRLTHLPKXENAFASAD-----NTKKFSLPKEGFCHNMLM---QQ 385
QY 234 IYPVNAISPHNIHTFATGSGDGFVNIWDPFNKRLCQFHR---YPTSIASLA-----ES 285
DB 386 KTLINAMAV-NEDGMVMTGGDNGSINFWDMKSHGFSFQSQETIIVQPSLSEAGIYAACYD 444
```

```
QY 286 NDGTTTLAIASS-----YVWHDDETHPE 308
DB 445 NTGSLVTCADKTIKWKEDENATPE 471
RESULT 15
YH92_CABEL STANDARD; PRT; 501 AA.
AC Q23256;
DC 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 54.5 kDa Trp-Asp repeats containing protein ZC302.2 in
DE Chromosome V.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Kelley P.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 7 WD repeats.
CC -!- SIMILARITY: STRONG, TO K04G11.4 AND C14B1.4.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z73978; CAA98293.1; --
CC FIRM; Z7513; T27513.
CC WormPep; ZC302.2; CE06574.
CC InterPro; IPR001680; WD40.
CC Pfam; PF00400; WD40; 7.
CC PRINTS; PR00320; GPROTEINBRPT.
CC ProDom; PD000018; WD40; 2.
CC SMART; SM00320; WD40; 7.
CC PROSITE; PS00678; WD_REPEATS_1; 3.
CC PROSITE; PS50082; WD_REPEATS_2; 6.
CC PROSITE; PS50294; WD_REPEATS_REGION; 1.
CC Hypothetical protein; Repeat; WD repeat.
CC REPEAT 211 241 WD 1.
CC REPEAT 253 283 WD 2.
CC REPEAT 295 325 WD 3.
CC REPEAT 337 367 WD 4.
CC REPEAT 381 410 WD 5.
CC REPEAT 422 455 WD 6.
CC REPEAT 467 499 WD 7.
CC SEQUENCE 501 AA; 54473 MW; DB22116996F85988 CRC64;
Query Match 9.8%; Score 173; DB 1; Length 501;
Best Local Similarity 25.6%; Pred. No. 2.1e-07;
Matches 69; Conservative 41; Mismatches 108; Indels 52; Gaps 13;
QY 16 GISVKFSPNTSQFLVSSWDTSVRLYVPANSMRLKYQHTGAVLPCAFYDPTHAW--S 72
DB 257 GINDFSSNSN-SQFIASASDVTVKIFDVISGACLRMTGRHTVYVCCSF-NPQSSLIAS 314
QY 73 GGLDQKMDLNTDQ-ENLVGTHDAPIRCVCEYCPVNVAVTGSWDTVKLMDPR- 127
DB 315 AGFDETYRVWDFKTGLCVKCPAHSDFPITSYNHGDNTWATSSYDGCIRVWDAAGSCL 374
QY 128 PCNAGTFSQEKVYVTLVSGDRLIVGTAGRRVLVWDLRNNGYVQQRRESSLKY-----Q 181
DB 375 KTLVDTTHAPVTVCFSPNGKYLISQALDSSLKLD-----PKAKPLKYNGHKNK 426
```

QY 182 TRCIRA---FENKQGYVLSIEGRVAVEYLDPSPEVQKKYAFKCHRLKENNIEQIY--- 235
Db 427 KYCLFANMSVPLGKHIISSGSEDGRILW-----SIQIKQ-----IVQILLEGH 468
QY 236 --PYN AISFHNHNTFATGG--SDGFVNIW 261
Db 469 TTPVLATDSHPTLNIISGGLPEPDNVIRIW 498

Search completed: November 12, 2003, 20:09:20
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 20:06:37 ; Search time 41 Seconds
(without alignments)
2064.421 Million cell updates/sec

Title: US-10-084-700-2
Perfect score: 1772
Sequence: 1 MTGSNEFKLNQPEDGISSV.....DGIFIRQVTDATKPKSPCT 328

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPREMBL 23.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertebrate.*
- 14: sp.unclassified.*
- 15: sp.virus.*
- 16: sp.bacteriaiap.*
- 17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1651	93.2	324	13 Q98UH2	Q98uh2 xenopus lae
2	1651	93.2	330	13 Q9YGY3	Q9ygy3 xenopus lae
3	1050.5	59.3	326	5 Q9VAJ2	Q9vaj2 drosophila
4	1035.5	58.4	326	5 Q9XYF7	Q9xyf7 drosophila
5	949	53.6	327	5 Q76523	O76523 drosophila
6	949	53.6	339	10 Q9C701	Q9c701 arabidopsis
7	939	53.0	340	10 Q9LJN8	Q9ljn8 arabidopsis
8	822.5	46.4	323	5 Q8MSW6	Q8msw6 drosophila
9	710	49.1	343	5 Q9XWH0	Q9xwh0 caenorhabdi
10	691	39.0	250	10 Q9C6C4	Q9c6c4 arabidopsis
11	671.5	37.9	357	3 Q59902	O59902 emaricella
12	586	33.1	344	10 Q8RYM1	Q8rym1 oryza sativ
13	578.5	32.6	349	10 Q9SAJ0	Q9saj0 arabidopsis
14	541.5	30.6	362	3 Q74224	O74224 emaricella
15	530.5	29.9	349	3 Q9HEB3	Q9heb3 neurospora
16	523.5	29.5	349	5 Q9GRX8	Q9grx8 chironomus

17	521.5	29.4	346	5 Q9W2E7	Q9w2e7 drosophila
18	518.5	29.3	320	3 Q42860	Q42860 schizosacch
19	517	29.2	368	11 Q8C570	Q8c570 mus musculus
20	508	28.7	341	11 Q8CBU3	Q8cbu3 mus musculus
21	492	27.8	208	5 Q8MPF0	Q8mpf0 taenia soli
22	457	25.8	314	10 Q8LA29	Q8la29 arabidopsis
23	452.5	25.5	315	10 Q9C795	Q9c795 arabidopsis
24	431	24.3	82	11 Q9CS16	Q9csi6 mus musculus
25	325	18.3	336	5 Q9W1T0	Q9wit0 drosophila
26	325	18.3	336	5 Q8T4A6	Q8t4a6 drosophila
27	316	17.8	318	5 Q8SRM6	Q8srme encephalito
28	272	15.3	339	5 Q25349	Q25349 leishmania
29	231	13.0	45	11 Q8BH42	Q8bh42 mus musculus
30	215	12.1	934	16 Q8Y2Z3	Q8y2z3 anabaena sp
31	206	11.6	1227	16 Q8Z0R1	Q8z0r1 anabaena sp
32	198.5	11.2	140	10 Q9LMQ0	Q9lmq0 arabidopsis
33	198	11.2	476	11 Q8BHD1	Q8bhd1 mus musculus
34	196	11.1	358	5 Q9W3J1	Q9w3j1 drosophila
35	194.5	11.0	1747	16 Q8Z020	Q8z020 anabaena sp
36	192	10.8	356	11 Q9CY09	Q9cy09 mus musculus
37	192	10.8	405	11 Q8J2X3	Q8j2x3 mus musculus
38	188	10.6	329	16 Q8KB12	Q8kb12 chlorobium
39	187	10.6	1189	16 Q8YL09	Q8yl09 anabaena sp
40	187	10.6	1551	16 Q8YMU3	Q8ymu3 anabaena sp
41	186	10.5	297	3 Q9C1X0	Q9c1x0 schizosacch
42	185.5	10.5	357	4 Q95320	Q95320 homo sapien
43	185.5	10.5	407	4 Q8NBT0	Q8nbt0 homo sapien
44	184.5	10.4	415	4 Q8N136	Q8n136 homo sapien
45	183.5	10.4	415	4 Q8N776	Q8n776 homo sapien

ALIGNMENTS

RESULT 1

Q98UH2 ID Q98UH2 PRELIMINARY; PRT; 324 AA.
AC Q98UH2;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE WD repeat protein Bub3.
GN BUB3.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI_TaxID=8355;

RN [1]_TaxID=8355;

SEQUENCE FROM N.A.

RP MEDLINE=21154833; PubMed=11231148;

RA Schwab M., Roberts B.T., Gross S.D., Tunquist B.J., Taieb F.E.,

RA Lewellyn A., Maller J.L.;

RT "Bub1 is activated by the protein kinase p90Rsk during Xenopus oocyte

RT maturation.";

RL Curr. Biol. 11:141-150(2001).

DR EMBL; AF119790; AAK12629.1; -

DR InterPro; IPR001680; WD40.

DR Pfam; PF00400; WD40. 5.

DR PRINTS; PF00320; GPROTEINRPT.

DR ProDom; PD000018; WD40; 1.

DR SMART; SM00320; WD40; 5.

DR PROSITE; PS50082; WD_REPEATS_2; 3.

DR PROSITE; PS50294; WD_REPEATS_REGION; 2.

DR Repeat; WD repeat.

SQ SEQUENCE 324 AA; 35562 MW; 417B7C744417F01B CRC64;

Query Match

Best Local Similarity 93.2%; Score 1651; DB 13; Length 324;

Matches 302; Conservative 14; Mismatches 8; Indels 0; Gaps 0;

Qy

1 MTGSNEFKLNQPEDGISSVFKFSPNTSQFLVSSWTSVRLYDVPANSMELKYQHTGAVL 60
|||||

```
Db 1 MTGSNEPKLNQAPEDGISAVKFSPTNTSOFLLVSSNDSSVRLYDVPANTMLRKYQHAGPVL 60
Qy 61 DCAFYDPTAHWSGGDLHQLKHDINTDQENLVGTHDAPIRCVVEYCPVNVNMTGSNDQTV 120
Db 61 DCAFYDPTAHWSGGDLHQLKHDINTDQENLVGTHDAPIRCVVEYCPVNVNMTGSNDQTV 120
Qy 121 KLMDPRTPCNAGTFSQPEKVTYLSVSGDRLLVGTAGRRVLVMDLRNMGVYVQORRESSLKY 180
Db 121 KLMDPRTPCNAGTFSQPEKVTYLSVSGDRLLVGTAGRRVLVMDLRNMGVYVQORRESSLKY 180
Qy 181 QTRCIRAFNPKQGYVLLSSIEGRVAVVEYLDPSVEVQKKYAPKCHRLKENNIEQIYPNAV 240
Db 181 QTRCIRAFNPKQGYVLLSSIEGRVAVVEYLDPSVEVQKKYAPKCHRLKENNIEQIYPNAV 240
Qy 241 SFPHNHTFATGSGDGFVNWDPPNKKRLCQFHRYPTSIASLAFSNDGTTLAIASSTMYE 300
Db 241 SFPHNHTFATGSGDGFVNWDPPNKKRLCQFHRYPTSIASLAFSNDGTTLAIASSTMYE 300
Qy 301 MDDTEHPEDGFIROVTDATKPK 324
Db 301 MDDIDHPEDAIYIROVTDATKPK 324

RESULT 2
QYGVY3
ID QYGVY3 PRELIMINARY; PRT; 330 AA.
AC QYGVY3
DT 01-MAY-1999 (TREMELrel. 10, Created)
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
DE Mitotic checkpoint.
GN XBUB3.
OS Xenopus laevis (african clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99196240; PubMed=10098604;
RA Goto T., Kinoshita T.;
RT "Maternal transcripts of mitotic checkpoint gene, Xbub3, are
RT accumulated in the animal blastomeres of Xenopus early embryo.";
RL DNA Cell Biol. 18:227-234 (1999).
DR EMBL; AB018419; BAA34999.1; -
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS50082; WD_REPEATS_2; 3.
DR PROSITE; PS50294; WD_REPEATS_REGION; 2.
KW Repeat; WD repeat.
SQ SEQUENCE 330 AA; 37253 NW; 8075DC94738F5E19 CRC64;

Query Watch 93.2%; Score 1651; DB 13; Length 330;
Best Local Similarity 93.2%; Pred. No. 2.9e-139;
Matches 302; Conservative 14; Mismatches 8; Indels 0; Gaps 0;

QY 1 MTGSNEPKLNQAPEDGISAVKFSPTNTSOFLLVSSNDSSVRLYDVPANMLRKYQHAGPVL 60
Db 7 MTGSNEPKLNQAPEDGISAVKFSPTNTSOFLLVSSNDSSVRLYDVPANMLRKYQHAGPVL 66
Qy 61 DCAFYDPTAHWSGGDLHQLKHDINTDQENLVGTHDAPIRCVVEYCPVNVNMTGSNDQTV 120
Db 67 DCAFYDPTAHWSGGDLHQLKHDINTDQENLVGTHDAPIRCVVEYCPVNVNMTGSNDQTV 126
Qy 121 KLMDPRTPCNAGTFSQPEKVTYLSVSGDRLLVGTAGRRVLVMDLRNMGVYVQORRESSLKY 180
Db 127 KLMDPRTPCNAGTFSQPEKVTYLSVSGDRLLVGTAGRRVLVMDLRNMGVYVQORRESSLKY 186
Qy 181 QTRCIRAFNPKQGYVLLSSIEGRVAVVEYLDPSVEVQKKYAPKCHRLKENNIEQIYPNAV 240
Db [1]
```

```
Db 187 QTRCIRAFNPKQGYVLLSSIEGRVAVVEYLDPSVEVQKKYAPKCHRLKENNIEQIYPNAV 246
Qy 241 SFPHNHTFATGSGDGFVNWDPPNKKRLCQFHRYPTSIASLAFSNDGTTLAIASSTMYE 300
Db 247 SFPHNHTFATGSGDGFVNWDPPNKKRLCQFHRYPTSIASLAFSNDGTTLAIASSTMYE 306
Qy 301 MDDTEHPEDGFIROVTDATKPK 324
Db 307 MDDIDHPEDAIYIROVTDATKPK 330

RESULT 3
QYVAJ2
ID QYVAJ2 PRELIMINARY; PRT; 326 AA.
AC QYVAJ2;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
DE BUB3 protein.
GN BUB3 OR CG7581.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKLEY;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champ M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Feiraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195 (2000).
DR EMBL; AB003770; AAF56914.1; -
DR FlyBase; FB00025457; BUB3.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINBRPT.
DR SMART; SM00320; WD40; 5.
```

```

DR PROSITE; PS00678; WD_REPEATS_1; 1.
DR PROSITE; PS50082; WD_REPEATS_2; 2.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 326 AA; 37415 MW; 89EC16E18ESA9E9F CRC64;

Query Match 59.3%; Score 1050.5; DB 5; Length 326;
Best Local Similarity 60.6%; Pred. No. 1.3e-85;
Matches 195; Conservative 49; Mismatches 75; Indels 3; Gaps 2;

QY 6 EFKLNQPEDGIGSVKSPNTSQFLVSSWMTSVRLYDVPANSMRLKYQHTGAVLDCAFY 65
DB 5 EFKLNPPEDLISAVKFGPKSNQYMAASSWDGTLRFYDVPANQLRQKQFVQDAPLLDCAFM 64

QY 66 DPTHAWSGLDHQLKMHDLNTDQENLVGTHDAPIRCVEYCFEYVNMVMTGSDQTVKLWDP 125
DB 65 DIVHVSGSLDQLRFLFDVNTQAESIIIGAHEEPIKCLEHAEEYVNGILTGSDQTVKLWDM 124

QY 126 RTPCNAGTFSQPE-KVYTLVSGDRLLVGTAGRRVLVWDLRNMGYVQORRESSLKQYQTRC 184
DB 125 REKRCVGTGFEQNGKGVYMSVIDEKIVVATSDRKLWDLRKMDSYIMKRESSLKQYQTRC 184

QY 185 IRAFPNKGYSVLSIEGRVAVVEYLDPSPEVQKKYAFKCHRLKENNIEQIYFPVNAISFHN 244
DB 185 IRLFPNKGYSVMSIEGRVAVVEYLDHDPVQRRKFAFKCHRNREQNIEQIYFPVNAISFHN 244

QY 245 IHNTPATGSGDGFVNINDPFPNKKRLCOFHPYPTSIASLAFSNDGTTLAIASSYMYEMDDT 304
DB 245 VIQFATGSGSGIYVINDGFPNKKRLCOFHEYDTSISTINFTSSDGSALAISSYLDQLPET 304

QY 305 EH--PEDGIFIRQVTDAPTKPK 324
DB 305 PATVHPAIIYIRYPTDQETQK 326

RESULT 5
QYXIF7 PRELIMINARY; PRT; 327 AA.
ID O76523
AC O76523;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Mitotic checkpoint control protein BUB3.
GN BUB3 OR CG7581.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Basu J., Logazinho E., Herrman S., Bousbaa H., Li Z., Chan G.K.T.,
RA Yen T.J., Sunkel C.E., Goldberg M.L.;
RT "Localization of the Drosophila checkpoint control protein BUB3 to the
RT kinetochore requires Bub1 but not Zw10 or Rod.";
RL Chromosome 0:0-0(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=99115384; PubMed=9914369;
RA Basu J., Logazinho E., Herrmann S., Bousbaa H., Li Z., Chan G.K.,
RA Yen T.J., Sunkel C.E., Goldberg M.L.;
RT "Localization of the Drosophila checkpoint control protein BUB3 to the
RT kinetochore requires Bub1 but not Zw10 or Rod.";
RL Chromosome 107:376-385(1998).
DR EMBL; AF075594; AAD13398.1; -.
DR EMBL; AF088846; AAD44035.1; -.
DR FlyBase; FBgn0025457; BUB3.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINBRPT.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
DR PROSITE; PS0082; WD_REPEATS_2; 2.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 327 AA; 37047 MW; 8205F42395852A30 CRC64;

Query Match 53.6%; Score 949; DB 5; Length 327;
Best Local Similarity 56.3%; Pred. No. 1.5e-76;
Matches 183; Conservative 50; Mismatches 84; Indels 8; Gaps 4;

QY 6 EFKLNQPEDGIGSVKSPNTSQFLVSSWMTSVRLYDVPANSMRLKYQHTGAVLDCAFY 65
DB 5 EFKLNPPEDLISAVKFGPKSNQYMAASSWDGTLRFYDVPANQLRQKQFVQDAPLLDCAFM 64

QY 66 DPTHAWSGLDHQLKMHDLNTDQENLVGTHDAPIRCVEYCFEYVNMVMTGSDQTVKLWDP 125

```



```

Query Match      53.0%; Score 939; DB 10; Length 340;
Best Local Similarity 54.1%; Pred. No. 1.3e-75;
Matches 178; Conservative 54; Mismatches 91; Indels 6; Gaps 2;

QY 1 MTGNEFKLQPPEDGIGSSVKFSPNTSQFLVSSWDTSLVRLYDVPANSMRLKYQHTGAVL 60
DB 4 VTPSAGRELNSPDSGSLNRLFS--NNSDHLVSSNDKRVLYDVSTNSLKGELHGAVAL 62

QY 61 DCAFYDTHAWSGGLDHLQKMDLNTQENLVGTHDAPIRCVCEYCEPVNVMVTGSDQTV 120
DB 63 DCCPHDDFSGFSGADYKVRIRIVFNVGKEDILGTHDKAVRCVCSYAAAGQVITGSDKTV 122

QY 121 KLWDPK-----TPCNAGTFSQPEKVTLSVSGDRILVGTAGRRVLVMDLRNMGVYQORRE 175
DB 123 KCWDPKASGPFRTQVGTLYQPERVYMSLVGHLRVVATAGRHVNIYDLNKSQPORRE 182

QY 176 SSLKYQTRCIAFPNKGQYVLSSTEGRAVEYLDPSPEVQKKYAFKCHRLKNNIEQY 235
DB 183 SSLKYQTRCVRCYENGRTGYALSSVEGRVAMEFFDLSAAQAKYAFKCHRKSEAGRDIVY 242

QY 236 PVNAISPHNTHNTATGSGDGFVNIPFNKRLCOFHYPTSIASLAFSNDGTTLAIAS 295
DB 243 PVNSIAHPHYGTATGCGDGFVNWDGNKKRLYQSKYPTSIASLSFRDQQLAVAS 302

QY 296 SYMYEMDTEHPEDGIFIRQVTDATKPK 324
DB 303 SYTEGEKSGPEAIFVRSNTEVRFK 331

RESULT 8
Q8MSW6 PRELIMINARY; PRT; 323 AA.
ID Q8MSW6
AC Q8MSW6;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE LD23540P.
GN BUB3 OR CG7581.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Megaloptera; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dzenek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Han K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY118526; AAM49895.1;
DR FlyBase; FBgn0025457; BUB3.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40.4
DR PRINTS; PR00320; GPROTEINRPT.
DR SMART; SMO0320; WD40.4;
DR PROSITE; PS00678; WD REPEATS 1; 1.
DR PROSITE; PS00082; WD REPEATS 2; 1.
DR PROSITE; PS50294; WD REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 323 AA; 37515 MW; 5AECF3AB8D5154B7 CRC64;

Query Match      46.4%; Score 822.5; DB 5; Length 323;
Best Local Similarity 60.7%; Pred. No. 3e-65;
Matches 150; Conservative 40; Mismatches 56; Indels 1; Gaps 1;

QY 6 EFKLNPEDGIGSSVKFSPNTSQFLVSSWDTSLVRLYDVPANSMRLKYQHTGAVLCAFY 65
DB 5 EFKLNPEDLISAVKFGPKSNQWAASSWDGTLRFDVPANQIRKQFVQADAPLLOCAF 64

us-10-084-700-2.rspt
Query Match      53.0%; Score 939; DB 10; Length 340;
Best Local Similarity 54.1%; Pred. No. 1.3e-75;
Matches 178; Conservative 54; Mismatches 91; Indels 6; Gaps 2;

QY 1 MTGNEFKLQPPEDGIGSSVKFSPNTSQFLVSSWDTSLVRLYDVPANSMRLKYQHTGAVL 60
DB 4 VTPSAGRELNSPDSGSLNRLFS--NNSDHLVSSNDKRVLYDVSTNSLKGELHGAVAL 62

QY 61 DCAFYDTHAWSGGLDHLQKMDLNTQENLVGTHDAPIRCVCEYCEPVNVMVTGSDQTV 120
DB 63 DCCPHDDFSGFSGADYKVRIRIVFNVGKEDILGTHDKAVRCVCSYAAAGQVITGSDKTV 122

QY 121 KLWDPK-----TPCNAGTFSQPEKVTLSVSGDRILVGTAGRRVLVMDLRNMGVYQORRE 175
DB 123 KCWDPKASGPFRTQVGTLYQPERVYMSLVGHLRVVATAGRHVNIYDLNKSQPORRE 182

QY 176 SSLKYQTRCIAFPNKGQYVLSSTEGRAVEYLDPSPEVQKKYAFKCHRLKNNIEQY 235
DB 183 SSLKYQTRCVRCYENGRTGYALSSVEGRVAMEFFDLSAAQAKYAFKCHRKSEAGRDIVY 242

QY 236 PVNAISPHNTHNTATGSGDGFVNIPFNKRLCOFHYPTSIASLAFSNDGTTLAIAS 295
DB 243 PVNSIAHPHYGTATGCGDGFVNWDGNKKRLYQSKYPTSIASLSFRDQQLAVAS 302

QY 296 SYMYEMDTEHPEDGIFIRQVTDATKPK 324
DB 303 SYTEGEKSGPEAIFVRSNTEVRFK 331

RESULT 8
Q8MSW6 PRELIMINARY; PRT; 323 AA.
ID Q8MSW6
AC Q8MSW6;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE LD23540P.
GN BUB3 OR CG7581.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Megaloptera; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dzenek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Han K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY118526; AAM49895.1;
DR FlyBase; FBgn0025457; BUB3.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40.4
DR PRINTS; PR00320; GPROTEINRPT.
DR SMART; SMO0320; WD40.4;
DR PROSITE; PS00678; WD REPEATS 1; 1.
DR PROSITE; PS00082; WD REPEATS 2; 1.
DR PROSITE; PS50294; WD REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 323 AA; 37515 MW; 5AECF3AB8D5154B7 CRC64;

Query Match      46.4%; Score 822.5; DB 5; Length 323;
Best Local Similarity 60.7%; Pred. No. 3e-65;
Matches 150; Conservative 40; Mismatches 56; Indels 1; Gaps 1;

QY 6 EFKLNPEDGIGSSVKFSPNTSQFLVSSWDTSLVRLYDVPANSMRLKYQHTGAVLCAFY 65
DB 5 EFKLNPEDLISAVKFGPKSNQWAASSWDGTLRFDVPANQIRKQFVQADAPLLOCAF 64

```

```

QY 66 DPTHAWSGGLDHLQKMDLNTQENLVGTHDAPIRCVCEYCEPVNVMVTGSDQTVKLWDP 125
DB 65 DIVHWVSGSDNQLRLOFQNTQRESIIIGAHEEPIRCVHEAEYNGIUTGSDWNTVKLWDM 124

QY 126 RTPCNAGTFSQPE-KVYTLTSLVSGDRILVGTAGRRVLVMDLRNMGVYQORRESSLK YQTRC 184
DB 125 REKRCVGTGFEQNGKQVSMVIDEKIVVATSDRKVLWDLRKWDSYIMKRESSLK YQTRC 184

QY 185 IRAPNPKQGYVLSSTEGRAVEYLDPSPEVQKKYAFKCHRLKNNIEQYIPVNAISFHN 244
DB 185 IRLFPNKEGYVMSSTEGRAVEYLDHDPPEVQRRKFAFKCHRNREQNIEQYIPVNAISFHN 244

QY 245 IHNTPAT 251
DB 245 VYQTFAT 251

RESULT 9
Q9XWH0 PRELIMINARY; PRT; 343 AA.
ID Q9XWH0
AC Q9XWH0;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
GN Y54G9A.6
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Smye R.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=99069613; PubMed=9851916;
DE none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AL032648; CAA21698.1;
DR WormPep; Y54G9A.6; CE19233;
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40.5.
DR PRINTS; PR00320; GPROTEINRPT.
DR SMART; SMO0320; WD40.3.
DR PROSITE; PS00678; WD REPEATS 1; 1.
DR PROSITE; PS00082; WD REPEATS 2; 2.
DR PROSITE; PS50294; WD REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 343 AA; 38127 MW; 021AD0398C6BDS5D CRC64;

Query Match      40.1%; Score 710; DB 5; Length 343;
Best Local Similarity 43.6%; Pred. No. 3.7e-55;
Matches 144; Conservative 67; Mismatches 109; Indels 10; Gaps 6;

QY 5 NEFKLNQPPEDGIGSSVKFSPNT-SQFLVSSWDTSLVRLYDVP-NSMRLKYQHTGAVL 60
DB 14 NEFRVPFPFPVQISKVQFORAGSRLLAASGMDGTCRVYEVGKLGDISEKLVTHGKPL 73

QY 61 DCAFYDTHAWSGGLDHLQKMDLNTQENLVGTHDAPIRCVCEYCEPVNVMVTGSDQTV 120
DB 74 TCTAGYKVAFGGVGDHNVKLADIETGCTQLGSHALAVRCMFPNMFSSLIIVSGWDSV 133

QY 121 KLWDPKCNAG--TFSQPEKVTLSVSGDRILVGTAGRRVLVMDLRNMGVYQORRESSL 178
DB 134 KLWDARSYNGAIESVNVSSVYAMDVLKHTILVGTDKRKIFMYDSRKREPLQVRDPL 193

QY 179 KYOTECIRAPNKGQYVLSSTEGRAVEYLDPS-PEVQKKYAFKCHRLKE-NNIEQIYP 236
DB 194 KYOTRAVOFFPTGFAFVSSIEGRVAVEYVDQSGEEMMKRYAFKCHREKDTGDTLHP 253

```


DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Putative mRNA export protein.
 GN P0663E10.24.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoidae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (CA3) genomic DNA, chromosome 1, PAC
 CLONE: P0663E10.24";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP004317; BAB90741.1; -;
 DR Gramene; O8RIN1; -;
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 4.
 DR SMART; SM00320; WD40; 5.
 DR PROSITE; PS00082; WD REPEATS 2; 4.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 SQ SEQUENCE 344 AA; 37849 MW; 9C65902197F6001EC CRC64;
 Query Match 33.1%; Score 586; DB 10; Length 344;
 Best Local Similarity 38.8%; Pred. No. 4.5e-44;
 Matches 128; Conservative 68; Mismatches 118; Indels 16; Gaps 9;
 QY 4 SNEFKLNQPPEDGISSVKFSPNTSQPLVSSWDTSVRLVDV-PANSMRLK--YQHTGAVL 60
 Db 16 NKLEVNPPAGDSVSSLSFPAKH-LVATSDNQRVCEIOPGGCCQAKASISHDQVPL 74
 QY 61 DCAFYDP-THAMSGGLDHLKMD-LNTQENLVGTHDAPIRVCYCEPVNVMVWGSDQ 118
 Db 75 CSAWKDDGTVFSGGDKQIKWPLLSGGQPMVLSGHEAPVKELAWIFQMLNLSGSDK 134
 QY 119 TVKLMDPRTCPNAGTSQPEKYVTLVSVDRLIVGTAGRLVWDLRWNGVVOQRRESL 178
 Db 135 TLRYDVRQQAHPVQOLPERCYALSPLVMVGTADRNVTFNQNQPAEFKRIIVSPL 194
 QY 179 KYQTRCIRAPPNKQGYVLSIEGRVAVELDPSPEVQKKYAFKCHRLKNNIEIYVNV 238
 Db 195 KLQTRCLAAFPDQGFVLSIEGRVAVELDPSPEVQKKYAFKCHRLKNNIEIYVNV 248
 QY 239 AISFNHNTFATGSDGFVNWDPNKRLCQPHRYPTSIASLAFSNDGTTLAASSYM 298
 Db 249 ALNFHPVHTFATGSDGFVNWDPNKRLCQPHRYPTSIASLAFSNDGTTLAASSYM 298
 QY 299 YEMDDTEH----PEDGIFIRQVTDATKPK 324
 Db 309 WSKGAENKPNSTAKTNIFLHVSQSEVCKK 338
 RESULT 13
 Q9SAJ0 PRELIMINARY; PRT; 349 AA.
 ID Q9SAJ0
 AC Q9SAJ0
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE F23A5.2(form2) (mRNA export protein, putative).
 GN F23A5.2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S., Lee J.M., Li J.,

RA Gonzalez A., Liu A., Liu K., Sakano H., Koo T., Pham P., Vaysberg M.,
 RA Hwang B., Chin C., Choi E., Chiou J., Altafi H., Brooks S., Chao Q.,
 RA Conn L., Conway A., Hansen N., Johnson-Hopson C., Khan S., Kim C.,
 RA Lam B., Nguyen M., Palm C., Shinn P., Tambunga G., Davis R.W.,
 RA Ecker J.R., Federpiel N.A., Theologis A.;
 RT "Arabidopsis thaliana chromosome 1 BAC F23A5 sequence";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Theologis A.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 annotation";
 RL Genome Biol. 0:0-0(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RT "Full-length cDNA from Arabidopsis thaliana";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC011713; AAP14655.1; -;
 DR EMBL: AY087683; AAM65220.1; -;
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 5.
 DR PRINTS; PR00320; GPROTEINRPT.
 DR PROSITE; PS00082; WD REPEATS 2; 4.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 SQ SEQUENCE 349 AA; 38268 MW; F1780EF3258C5632 CRC64;
 Query Match 32.6%; Score 578.5; DB 10; Length 349;
 Best Local Similarity 38.7%; Pred. No. 2.2e-43;
 Matches 129; Conservative 60; Mismatches 125; Indels 19; Gaps 8;
 QY 4 SNEFKLNQPPEDGISSVKFSPNTSQPLVSSWDTSVRLVDV-----ANSRLKYQHTG 57
 Db 15 NKSEVTEPSPADSISLSESPR-ADILNATSDNQRVCEIERSGASLASAPKASISHDQ 73
 QY 58 AVLDCAFYDP-THAMSGGLDHLKMD-LNTQENLVGTHDAPIRVCYCEPVNVMVWGSDQ 115
 Db 74 FVLCSAWKDDGTVFSGGCDKQAKMPLLSGGQPMVLSGHEAPVKELAWIFQMLNLSGSDK 133
 QY 116 WDQTVKLMDPRTCPNAGTSQPEKYVTLVSVDRLIVGTAGRLVWDLRWNGVVOQRRE 175
 Db 134 WDKTLKYWDTRQNPVHTQOLPKCVTLVSKHPLVMVGTADRNVTFNQNQPAEFKRIQ 193
 QY 176 SSLKYQTRCIRAPPNKQGYVLSIEGRVAVELDPSPEVQKKYAFKCHRLKNNIEIY 235
 Db 194 SPLKYQTRCIRAPPNKQGYVLSIEGRVAVELDPSPEVQKKYAFKCHRLKNNIEIY 247
 QY 236 PVNAISFNHNTFATGSDGFVNWDPNKRLCQPHRYPTSIASLAFSNDGTTLAASSYM 295
 Db 248 SVNSLNFHPVHTFATGSDGFVNWDPNKRLCQPHRYPTSIASLAFSNDGTTLAASSYM 307
 QY 296 SYWYEMDDTEH----PEDGIFIRQVTDATKPK 324
 Db 308 CYDWSKGAENHPNATAPAKSSIFLHLPQSEVCKK 340
 RESULT 14
 O74224 PRELIMINARY; PRT; 362 AA.
 ID O74224
 AC O74224;
 DT 01-NOV-1998 (T-EMBLrel. 08, Created)
 DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE SONA.
 GN SONA.

OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=162425;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98311642; PubMed=9647650;
 RA Wu L., Osmani S.A., Mirabito P.M.;
 RT "A role for NIMA in the nuclear localization of cyclin B in
 aspergillus nidulans.";
 RL J. Cell Biol. 141:1575-1587(1998).
 DR EMBL; AF069492; AN027297.1; -
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 5.
 DR PRINTS; PR00320; GP0TEINBRPT.
 DR SMART; SM00320; WD40; 3.
 DR PROSITE; PS50082; WD REPEATS 2; 2.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 SQ SEQUENCE 362 AA; 39697 MW; 1E2BA08BFFED101E CRC64;
 Query Match 30.6%; Score 541.5; DB 3; Length 362;
 Best Local Similarity 36.2%; Pred. No. 4.6e-40;
 Matches 122; Conservative 53; Mismatches 143; Indels 19; Gaps 8;
 QY 4 SNEFKLNQPPEDGSSVKEFSPNTS--QFLVSSNDTSVRLYDVPANSM---RLKYQHTGAVL 60
 Db 23 SKDVALNSPPEDSISDURFSP-ASEHLAVASWDKKYRIEINQGSSEGRALFEHAPVL 81
 QY 61 DCAPY-DPTHAWSGGLDHLQKMDL--NTDQENLVGTHDAPIRCVCEYCPV---VMVMT 113
 Db 82 NCCVSPDQTKVVGAGADKAARMULDAAATAPIQVAADAPIRCCHMIENPAGTELLVT 141
 QY 114 GSWDQTKLVMDPRTPCNAFTSQPEKYVTLVSGDRLIVCTAGRVLVWDLNMGVYQOR 173
 Db 142 GSWDKQVYVWLRSTPIASLEQSERVYTMVDVKNLLVIGTADRYINIVDLNNTKFKYKT 201
 QY 174 RESLKYQTCIRAFPNKQGVVLSIEGRVAVEXYLDPSPEVQKKYAFKCHR---LKNN 230
 Db 202 MQSPFKWQTRVSVSCFTDANGFAVSGIEGRCAIQYVEDKD--SSNFSFKCHRETPNQRD 259
 QY 231 IEQIYPVNAISFHNHNTFTATGSDGFVNWDPNKRLQCFHYPTSTIASLAFSNDGTT 290
 Db 260 INNIVNAISFHFPHGTFTAGADGTFHFWDKAKHRLKGYPNVGGSTSTTAFNRTGNI 319
 QY 291 LAIASSYMYEMD---DTEHPEDGIFIRQVTDATKPK 324
 Db 320 FAYAVSYDWSKSGYSANTFQLPNKMVHLHPVAQEVEKPR 356

RESULT 15

Q9HEB3 ID Q9HEB3 PRELIMINARY; PRT; 349 AA.
 AC Q9HEB3;
 DT 01-MAR-2001 (TrEMBLrel. 15, Created)
 DT 01-MAR-2001 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Probable SONA.
 GN B11A5.110.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hoheisel J., Brandt P., Partmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL451109; CAC18615.1; -

DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 5.
 DR SMART; SM00320; WD40; 3.
 DR PROSITE; PS50082; WD_REPEATS 2; 2.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 2.
 KW Repeat; WD repeat.
 SQ SEQUENCE 349 AA; 38197 MW; E42D4C739B2430D8 CRC64;

Query Match 29.9%; Score 530.5; DB 3; Length 349;
 Best Local Similarity 37.5%; Pred. No. 4.2e-39;
 Matches 124; Conservative 54; Mismatches 132; Indels 21; Gaps 9;
 QY 6 EFKLNQPPEDGSSVKEFSPNTS---QFLVSSNDTSVRLYDVPANSMRLKYQHTG-AVLD 61
 Db 22 DVELGQPPEDSISDLAENPNPADQKFLAVASWDKTRIRYIELNSG-----QSGGQAMIE 76
 QY 62 CAFYDPTHAWSGGLDHLQKMDLNTDQENLVGTHDAPIRCVCEYCPV---VMVTCSDQ 118
 Db 77 HDAPDGTQVISAADKAALKVLDLATGSMQVAADMPKCVRYF-EANGTPMAVTTGWDK 135
 QY 119 TVKIWDPRTPCNAFTSQPEKYVTLVSGDRLIVCTAGRVLVWDLNMGVYQORRESSL 178
 Db 136 QIKYWDPRSANPAATVQAQERVYTMVDVKNLLVIGTADRYINIVLNKOPGKFKYTKMQSPL 195
 QY 179 KYQTRCIRAFPNKQGVVLSIEGRVAVEXYLDPSPEVQKKYAFKCHR-LKENNIEO1YPV 237
 Db 196 KMQTRVSVSCFNDSQGFAGISIEGRCAIQYVEDKDSA--SNFSFKCHRDPAQGNHTAVHAV 253
 QY 238 NAISFHNHNTFTATGSDGFVNWDPNKRLQCFHYPTSTIASLAFSNDGTTLAISSY 297
 Db 254 NDISFHPQHGTFSTAGSDGTFHFWDKAKHRLKGYPNVGGSTSTTTFNKTGSI FAYAISY 313
 QY 298 ----MYEMDTEHPEDGIFIRQVTDATKPK 324
 Db 314 DWSKGYGNSFTYPTK-VMLHPVQQDECKPR 343

Search completed: November 12, 2003, 20:10:14
 Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 20:07:13 ; Search time 20 seconds
(without alignments)
1577.166 Million cell updates/sec

Title: US-10-084-700-2

Perfect score: 1772

Sequence: 1 MTGSNFKLNQPPEDGISSV.....DGIFIRQVTDATKPKSPCT 328

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96169682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:**

1: pir1:**

2: pir2:**

3: pir3:**

4: pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	939	53.0	340	2 T52386	mitotic checkpoint
2	710	40.1	343	2 T27185	hypothetical prote
3	578.5	32.6	349	2 A96839	F23A5.2 (form2) (im
4	538	30.4	365	2 S0610	hypothetical prote
5	518.5	29.3	320	2 T38301	probable mitotic c
6	511.5	28.9	352	2 A56119	RNA export protein
7	418.5	23.6	373	2 T20723	hypothetical prote
8	356	20.1	341	1 B39554	cell cycle arrest
9	352.5	19.9	251	2 S74241	probable RNA expor
10	272	15.3	339	2 T02818	PolyA export prote
11	228	12.9	1526	2 AC2239	WD-40 repeat prote
12	215	12.1	934	2 A91889	WD-40 repeat prote
13	209	11.8	1258	2 A13155	WD-repeat protein
14	206	11.6	1227	2 AE1810	WD-40 repeat prote
15	198.5	11.2	140	2 H86292	F7H2.18 protein -
16	194.5	11.0	1747	2 AC1842	WD-40 repeat prote
17	187	10.6	1189	2 A12493	WD-repeat protein
18	187	10.6	1551	2 AB2410	WD-repeat protein
19	181	10.2	304	2 A91837	WD-40 repeat prote
20	178.5	10.1	1189	2 AH2154	WD-repeat protein
21	177.5	10.0	356	2 T22478	hypothetical prote
22	176.5	10.0	677	2 AE1861	serine/threonine k
23	176	9.9	1356	2 T18521	beta transducin-li
24	175.5	9.9	486	2 S49820	PRL1 protein - Ara
25	173	9.8	501	2 T27513	hypothetical prote
26	171.5	9.7	961	2 E86245	hypothetical prote
27	169.5	9.6	340	1 RGH01	GTP-binding regula
28	169.5	9.6	340	1 RGH01	GTP-binding regula
29	169.5	9.6	340	2 JC5057	G protein beta 1 -

30	169	9.5	431	2 S49821	PRL2 protein - Ara
31	168.5	9.5	642	2 T39490	transcription init
32	167.5	9.5	317	2 T46032	WD-40 repeat regul
33	167.5	9.5	1711	2 AD1842	WD-40 repeat prote
34	167	9.4	455	2 S61159	CDC40 protein - ye
35	166	9.4	318	2 S11904	GTP-binding regula
36	165.5	9.3	714	2 S56893	hypothetical prote
37	165	9.3	676	2 AH2195	hypothetical prote
38	164.5	9.3	614	2 S58306	WD-40 repeat regul
39	164	9.3	502	2 T41148	trp-asp repeat con
40	163.5	9.2	559	2 AB2202	hypothetical prote
41	163.5	9.2	1049	2 T42045	beta transducin-li
42	163.5	9.2	1683	2 AF2071	WD-40 repeat prote
43	162.5	9.2	326	2 T16587	GTP-binding protei
44	162.5	9.2	328	2 T16970	GTP-binding protei
45	161.5	9.1	326	2 T02340	GTP-binding regula

ALIGNMENTS

RESULT 1

T52386

mitotic checkpoint protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cross)

C;Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 24-Oct-2000

C;Accession: T52386

R;Kaneko, T.; Kato, T.; Sato, S.; Nakamura, Y.; Asamizu, E.; Tabata, S.

submitted to the EMBL Data Library, September 1999

A;Reference number: Z26062

A;Accession: T52386

A;Status: preliminary; translated from GS/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-340 <KAN>

A;Cross-references: EMBL:AP000417; PIDN:BA802543.1

A;Experimental source: cultivar Columbia

C;Genetics:

A;Map position: 3

A;Introns: 39/3; 83/2; 111/1; 200/1; 225/2; 252/2; 278/3; 312/2

Query Match

Best Local Similarity 53.0%; Score 939; DB 2; Length 340;

Matches 178; Conservative 54; Mismatches 91; Indels 6; Gaps 2;

Qy	1	MTGSNFKLNQPPEDGISSVFKSPMTSQFLVSSWDTSVRLVYDVFPANGMLKYQHTGAVL	60
Db	4	VTPSAGRELSPNPPSDGISNLRFS-NNSDHLVSSWMDKRVLYDVSTNSLKGFLHGGAVL	62
Qy	61	DCAFYDPTHAWSGGLDHLKMDLNTDQENLVGTHDAPIRCVCYCEPVNVMVWYTGSDQTV	120
Db	63	DCGFHDGFGSGVAGDYKVRIVFNVGKREDILGTHDKAVRCVEYSVAGQVITGSNDKTV	122
Qy	121	KLMDPR-----TPCNAQTSPQEKVYTLTVSGDRILVGTAGRRVLVWDLNNMGVYQORRE	175
Db	123	KCMDPRGASGPRTQGVTYLQPERVYSMSLVGHRLVWATAGRHVNIYDLRNNMSQEQRRRE	182
Qy	176	SSLYKQTRCIRAPNKGQVLSISGRVAVEYLDSPQVKKYAFKCHRLKENNIEQIY	235
Db	183	SSLYKQTRCVRCPVNTGTALSSVEGRVAMEFFDSEAAQAQKAYFKCHRKSEAGRIDVY	242
Qy	236	PVNAISFPHNHTTATGSGDGFVNWDPNKRLQCFHRYPTSIASLAFSNDGTTLAIS	295
Db	243	PVNSIAFHFIYGTATGCGDFVNWDGNKKRLQYKYPTISIALSFSRDLQLAVAS	302
Qy	296	SYNYEMDTEHPEDGIFIRQVTDATKPK	324
Db	303	SYTFEEGEKQBPFAIFRSVNEIEVKPK	331

RESULT 2

T27185

hypothetical protein Y54G9A.6 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 26-May-2000
 C:Accession: T27185
 R:Smyle, R.
 submitted to the EMBL Data Library, October 1998
 A:Reference number: Z20324
 A:Accession: T27185
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-343 <WLL>
 A:Cross-references: EMBL:AL032648; PIDN:CAA21698.1; GSPDB:GN00020; CRSP:Y54G9A.6
 A:Experimental source: clone Y54G9A
 C:Genetics:
 A:Gene: CESP:Y54G9A.6
 A:Map position: 2
 A:Introns: 78/3; 291/2
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 40.18; Score 710; DB 2; Length 343;
 Best Local Similarity 43.64; Pred. No. 3.3e-53;
 Matches 144; Conservative 67; Mismatches 109; Indels 10; Gaps 6;
 QY 5 NEFKLNQPPEDGISSVKSFPNT-SQFLVSSWDTSVRLYDVP-NSMRLKYQHTGAVL 60
 DB 14 NEFRVPPFPVQISKVQFQREAGSRLAASGMDGTCRVYEVGKLGDISKLVTHGKPL 73
 QY 61 DCAFYDPTHANGSGLDHOLKMDLNTQENLVGTHDAPIRCVCEVAVNMTGSDQTV 120
 DB 74 TCTFAGYKNAVAGVDHNVKLADYETNGTQLGSHALAVRCMFNMSLTVSGGWDSSV 133
 QY 121 KLMDPRTPCNAG--TFSPQPKVYTLVSGLDRLVGTAGRRVLVMDLRNMGVYQQRRESSL 178
 DB 134 KLWDARSVNGAIESVNVSSVYAMDVLKHTILVGTGKRIKIFWYDSKRLREPLQVDSPL 193
 QY 179 KYQTRCTRAPNKGQVLSIEGRVAVYLDPS-PEVOKKYAFKCHRLKE-NNTEQIYP 236
 DB 194 KYQTRAVQFFPTGAFVYSSIEGRVAVYLDPS-PEVOKKYAFKCHRLKE-NNTEQIYP 253
 QY 237 VNATSFNHNHTFATGSDGFVNIWDPFNKRLCQFHRYPTSIASLAFSNDGTTLAIASS 296
 DB 254 VHTVAFHPKYCTFATGAGDGIWNIWDPFNKRLCQFHRYPTSIASLAFSNDGTTLAIASS 313
 QY 297 YMY--EMDTEHPDGFIFIRQVDTAETPK 324
 DB 314 YOYEKIDPSPLPNNSTIRHITDPSRPK 343

RESULT 3
 A:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
 C:Accession: A96839
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 ansen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: A96839
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-349 <STO>
 A:Cross-references: GB:A8605173; NID:g6503279; PIDN:AAF14655.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F23A5.2
 A:Map position: 1
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 32.64; Score 578.5; DB 2; Length 349;
 Best Local Similarity 38.74; Pred. No. 6.5e-42;
 Matches 129; Conservative 60; Mismatches 125; Indels 19; Gaps 8;
 QY 4 SNEFKLNQPPEDGISSVKSFPNTSQFLVSSWDTSVRLYDVP-----ANSMRLKYQHTG 57
 DB 15 NKSEYVTPSPAD6ISLSFSPR-ADILVATSDNQCRCWEISRSKASLASAPKASISHDQ 73
 QY 58 AVLDCAFYDP-THANGSGLDHOLKMDLNTQENLVGTHDAPIRCVCEVAVNMTGVS 115
 DB 74 FVLCSAWKDDGTTFVSGGCDKQAKWPLLSGGQPTVAVHSGPTAAMAWIIFGMILLATGS 133
 QY 116 WDQTVKLDWPTPCNAGTFSQPEKVTYTLVSQDRLIVGTAGRRVLVMDLRNMGVYQQRRE 175
 DB 134 WDKTLKYMDTRQNPVHTQQLPKCYTSLSVKHLPMVVGTDNRNLIVFNLQNPQTEFRRIQ 193
 QY 176 SSKLYQTRCTRAPNKGQVLSIEGRVAVYLDPS-PEVOKKYAFKCHRLKENNIEQIY 235
 DB 194 SPUKYQTRCTRAPNKGQVLSIEGRVAVYLDPS-PEVOKKYAFKCHRLKENNIEQIY 247
 QY 236 PVNAISFNHNHTFATGSDGFVNIWDPFNKRLCQFHRYPTSIASLAFSNDGTTLAIAS 295
 DB 248 SVNSLNFHPVHGTFATGSDGAFNWDKSKQRLKAMSKCQPIPCSFNHDGSIYAVAA 307
 QY 296 SYNYEMDDTEH----PEDGIFIRQVDTAETPK 324
 DB 308 CYDWSKAGNHNPNATKSSIFLHLPQSEVRAK 340

RESULT 4
 S50610
 hypothetical protein YER107c - yeast (Saccharomyces cerevisiae)
 C:Species: Saccharomyces cerevisiae
 C:Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 23-Mar-2001
 C:Accession: S50610
 R:Dietrich, F.S.
 submitted to the EMBL Data Library, December 1994
 A:Description: The sequence of S. cerevisiae cosmid 9747, 8198, 9781, and lambda clone
 A:Reference number: S50436
 A:Accession: S50610
 A:Molecule type: DNA
 A:Residues: 1-365 <DIE>
 A:Cross-references: EMBL:U18839; NID:g603313; PIDN:AAB64662.1; PID:g603345; MIPS:YER107
 C:Genetics:
 A:Gene: SGD:GLE2
 A:Cross-references: SGD:S0000909; MIPS:YER107c
 A:Map position: 5K
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology
 F:32-65/Domain: WD repeat homology <WD1>
 F:76-109/Domain: WD repeat homology <WD2>

Query Match 30.44; Score 538; DB 2; Length 365;
 Best Local Similarity 35.74; Pred. No. 2.1e-38;
 Matches 122; Conservative 59; Mismatches 135; Indels 26; Gaps 10;
 QY 4 SNEFKLNQPPEDGISSVKSFPNTSQFLVSSWDTSVRLYDVPANSMRLKYQH-TGAVLDC 62
 DB 26 ANDIVINSPEADISIDIAFSPQDPMFSSASSWDGKVRWDVQVQVQRAQHSSSPVLC 85
 QY 63 AFY--DPTHANGSGLDHOLKMDLNTQENLVGTHDAPIRCVCEV--CPEVNV--MVTGSW 116
 DB 86 TRNSDGTKVASGGCDNALVDIASGGTQIQGHMSABIKVLRVQCCGFSNTECIVTGSW 145
 QY 117 DQTVKLDWPTPCNAGTFSQPEKVTYTLVSQDRLIVGTAGRRVLVMDLRNMGVYQQRRES 176
 DB 146 DKTIKVDNRQEQPTVSVNMPERVSMONKQSLVVAATAEHAIINJANPTTFKATTS 205
 QY 177 SLKYQTRCTRAPNKGQVLSIEGRVAVYLDPS-PEVOKKYAFKCHRLK-----E 228
 DB 206 PLKWQTRCVACNEADGVAIGSVGECRSIRYIDG--MOKSGFSFKCHRCRTNENRAPS 263
 QY 229 NNTEQIYPVNAISFNHNHTFATGSDGFVNIWDPFNKRLCQFHRYPTSIASL---AFS 285

Db 264 NGQSLVVPVNSIAFHPLYGTFVAGDGTGFNFWKQNRHL---KGYPTLQASIPVCSFN 320

Qy 286 NCGTTTALAIASSYMY---ENDDTEHEDGIFIRQVTTDAETKPK 324

Db 321 RNCGSVFAYALSYDMHQHMGWGNRPDPNVIRLHATTDDEVKEK 362

RESULT 5

T38301

probable mitotic checkpoint WD repeat protein - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000

R;Skellton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.; Wood, V.

submitted to the EMBL Data Library, September 1997

A;Reference number: T38301

A;Accession: T38301

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-320 <8KE>

A;Cross-references: EMBL:Z99163; PIDN:CA15243.1; GSPDB:GN00066; SPDB:SPAC23H3.08C

A;Experimental source: strain 972h; cosmid c23H3

C;Genetics:

A;Gene: SPDB:SPAC23H3.08C

A;Map position: 1

A;Introns: 108/1

C;Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 29.3%; Score 518.5; DB 2; Length 320;

Best Local Similarity 36.1%; Pred. No. 8.2e-37; Indels 21; Gaps 4;

Matches 115; Conservative 52; Mismatches 131;

14 EDGISSVKPEPNTSQELLVSSMDTSVRLYDVDPANSMRL-KYQHTGAVLDCFAFYPTTHAMS 72

12 KQIISVVFSPVSKNELIAGCGDGLHYQISENPPELLGKYLSSPILSLVETDEKALV 71

Qy 73 GGLDHLQKMDLNTDQENLVGTHDAPIRVCVEPVNVMVWGSMDQTVKLDNPTPCNAG 132

Db 72 GNLDTGVTTLDLNTRNHEFLGNHKGKVCISKLRLNCFISGWDKSRVWDVRY----- 126

Qy 133 TFSOP-----BKVYTLVSGDRLVAGTAGRRVLVMDLRNMGVYQQRRESSLKYQTRCI 185

Db 127 --KQVVGQDIQKIPASSSRDNLVLGCSERENLVYDIRNLKLPFQRRSPSKYMTSRV 184

Qy 186 RAFPENKQCYVLSIEGRVAVELDPSPVEQKKYAFKCHRLKENNIEQIYVNAISFNI 245

Db 185 CCNQNFEGFVSSIEGKTSVEYINPSQEAQSKNFTFKCHRQIQKDYDIYVYVNDLKHPFI 244

Qy 246 HNTFATGSGDPVNIWDPFNKRLCQPHRYPTSIASLAFSNDGTTTLAIASSYMYEMDDTE 305

Db 245 HQLATAGGQGVVAFWDIQVRKRLVLPNFSKINISSIFNVDSGLMATAICQA-----E 298

Qy 306 HPEDGIFIRQVTTDAETKPK 324

Db 299 EAAGNIYVHALESNFAAPK 317

RESULT 6

A56119

RNA export protein rael - fission yeast (Schizosaccharomyces pombe)

N;Alternate names: poly(a)+ rna export protein

C;Species: Schizosaccharomyces pombe

C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 26-May-2000

C;Accession: A56119; T39547

R;Brown, J.A.; Bharathi, A.; Ghosh, A.; Whalen, W.; Fitzgerald, E.; Dhar, R.

J. Biol. Chem. 270, 7411-7419, 1995

A;Title: A mutation in the Schizosaccharomyces pombe rael gene causes defects in poly(A)

A;Reference number: A56119

A;Accession: A56119

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-352 <BRO>

A;Cross-references: GB:U14951; NID:G625093; PIDN:AAA8631.1; PID:G625094

R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.

submitted to the EMBL Data Library, February 1998

A;Reference number: Z21862

A;Accession: T39547

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-352 <WOO>

A;Cross-references: EMBL:AL021748; PIDN:CA16856.1; GSPDB:GN00067; SPDB:SPEC16A3.05C

A;Experimental source: strain 972h; cosmid c16A3

C;Genetics:

A;Gene: rael

A;Map position: 2

A;Introns: 24/3; 50/2

C;Superfamily: unassigned WD repeat proteins; WD repeat homology

F;70-103/Domain: WD repeat homology <WD1>

Query Match 28.9%; Score 511.5; DB 2; Length 352;

Best Local Similarity 34.8%; Pred. No. 3.7e-36;

Matches 115; Conservative 58; Mismatches 142; Indels 15; Gaps 7;

Qy 6 EFKLNQPPEDGISSVKESPNTSQELLVSSMDTSVRLYDVDPANSM---RLKYQHTGAVLDC 62

Db 22 DVEVAQPPEDSISDLAFSPQ-AEYLAASSWDSKVRIVYEVQATGOSIGKALYHGGPVLSV 80

Qy 63 AF-YDPTHAWSGGLDHLQKMDLNTDQENLVGTHDAPIRVCVEYCPYV---VMVTSQMDQ 118

Db 81 NNSRDGTRKVASGSVDKSAKFEDIQTGQVQAADDAVRCVRFVEAMGTSPILATGSMOK 140

Qy 119 TVKLWDPRTPCNAGTFPSQPEKVTYLSVGDRLVITAGRRVLVMDLRNMGVYQQRRESSL 178

Db 141 TLKYWDLRQSTPIATVSLPERVYAMDCVHPPLTVATARNICVINGTSEPTKIFKLAMSP 200

Qy 179 KYQTCIRAFENKQCYVLSIEGRVAVELDPSPVEQKKYAFKCHRLKENNIEQIYVNA 238

Db 201 KQTRSLACFIKGDGYAIGSVGEGRCAIQNIDKNAQ--NFSFCHRNQAGNSADVYSVN 258

Qy 239 AISEFNHNTFATGSDGFVNIWDPFNKRLCQPHRYPTSIASLAFSNDGTTTLAIASSY- 297

Db 259 STAFHPQYGTFTSTAGSDGTFSDKQDSHQRLKSYPNVGGTISCTFTARTGDFAYAISYD 318

Qy 298 --MYEMDDTEHPEDGIFIRQVTTDAETKPK 324

Db 319 WSKGYTFENNAQLP-NKIMLHPVPQDEIKPR 347

RESULT 7

T20723

hypothetical protein F108.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 26-May-2000

C;Accession: T20723

R;Basham, V.

submitted to the EMBL Data Library, September 1996

A;Reference number: Z19315

A;Accession: T20723

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-373 <WIL>

A;Cross-references: EMBL:Z80216; PIDN:CA802280.1; GSPDB:GN00019; CESP:F10G8.3

A;Experimental source: clone F10G8

C;Genetics:

A;Gene: CESP:F10G8.3

A;Map position: 1

A;Introns: 36/3; 67/3; 345/3

C;Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 23.6%; Score 418.5; DB 2; Length 373;

Best Local Similarity 31.5%; Pred. No. 3.8e-28;

Matches 109; Conservative 63; Mismatches 145; Indels 29; Gaps 11;

Qy 2 TGSNEFKLNQPPEDGISSVKESPNTSQ--FLLVGSDWTSVRLYDV-PANSMRLKYQHT-- 56

Db 30 TQNDPLVDCGAPEDTIQVIKESPTPQDKPMLACGSDGTIRVWMDANTFEKQAQNIIP 89

QY 57 GAVLDCAFY-DPTHAMSGGLDHLQKMHDLNTOENLVGTHDAPIRCVYCEVFN-----V 110
DB 90 APILDIAWEDSSKIFACADKEARLWDLASQAVVGHDFGVKT---CHWINGNNVQC 146
QY 111 NVTGSMDQTVKLMPTPCNAGTFQ---PEKVYTLVSGDRLIVGTAGRRVLDLNNM 167
DB 147 LNTSGPDKTLRFDMKMLPNOTMAQIQLPFRVYADLVLPMAVALANKHIVYUENG 206
QY 168 GYVQORRESLKVQTCIRAFPNKQ-----GYVLSIEGRVAVELDPSPVQKKYAFK 222
DB 207 PTEVKDIESLQKQIRICISIFKQKSNQNPAGFALSGIEGRVAVQVYDVAN--PKDNFTFK 264
QY 223 CHRLKE--NNIEQIYPVNAISPHNIHTFATGSDGDFVNIWDFPNKRLCOFHRYPTISIA 280
DB 265 CHRSALVNGPQEIYAVNDICFHPQHGTLVTIGSDGSGYVSMWMDKARTKLTSEPHMELT 324
QY 281 SLAFSNDGTLTATASSYMEWDETHPEDG---IFIRQVTDATKPE 323
DB 325 CCDVHSSGAFVLYALGYDWSRHEGNTQPGSKIVIHKCIEDMKPRP 370

RESULT 8
B39654
cell cycle arrest protein BUB3 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein O2654; protein YOR026W
C:Species: Saccharomyces cerevisiae
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 21-Jul-2000
A:Accession: B39654; S54632; S66892; S19036
R:HOYT, M.A.; Totis, L.; Roberts, B.T.
Cell 66, 507-517, 1991
A:Title: Saccharomyces cerevisiae genes required for cell cycle arrest in response to 1c
A:Reference number: A39654; MUID:91330299; PMID:1651171
A:Accession: B39654
A:Molecule type: DNA
A:Residues: 1-341 <HOY>
A:Cross-references: GB:M64707; NID:g171136; PIDN:AAA4459.1; PID:g171137
R:de Haan, M.; Maarse, A.C.; Grivell, L.A.
submitted to the EMBL Data Library, May 1995
A:Reference number: S54617
A:Accession: S54632
A:Molecule type: DNA
A:Residues: 1-341 <DEH>
A:Cross-references: EMBL:X87331; NID:g1041652; PIDN:CNA60742.1; PID:g829136
A:Experimental source: strain S288C
R:de Haan, M.; Grivell, L.A.; Maarse, A.C.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S66877
A:Accession: S66892
A:Molecule type: DNA
A:Residues: 1-341 <DEW>
A:Cross-references: EMBL:Z74934; NID:g1420136; PIDN:CNA99216.1; PID:g1420137; GSPDB:GN00
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:BUB3; MIPS:YOR026W
A:Cross-references: SGB:S0005552; MIPS:YOR026W
A:Map position: 15R
C:Function:
A:Description: cell cycle arrest
C:Superfamily: cell cycle arrest protein BUB3
C:Keywords: cell cycle control

Query Match 20.1%; Score 356; DB 1; Length 341;
Best Local Similarity 30.1%; Pred. No. 7.9e-23;
Matches 103; Conservative 57; Mismatches 106; Indels 76; Gaps 15;
QY 8 KINQPPEDGISSVSKFSPNTSOFLLVSSWSDTSVRLY--DVPANSM---RLKYQHTGAVLD 61
DB 5 QTEQPKYISDKIIPKSKS--LLATISWDSGLTYKFDIAQKXVLLQSLRYKH--PLLC 61
QY 62 CAFYDPTHAMSGGLDHLQKMHDLNTOENLVGTHDAPIRCVYCEVFN-----PEVNV- 110
DB 62 CNFID-----NTDLQIVYGVTVQGSILKVDLIGSPFSQALTNNEANLG 103

QY 111 -----WVTGSMDQTVKLMPTPCNAGTFQ---PEKVYTLVSGDRLIVGTAGRRVLDLNNM 150
DB 104 ICRICKYGGDLKLIASMDGLIEVDPNRYGDVIAVKNLNNKTKYKNKIFMTDNTSSAL 163
QY 151 IVGTAGRRVLDL-----RNMGYVQORRESLKVQTCIRAFPNKQ-GYVLSIEGRVAV 205
DB 164 IVGMNNSQVQWFRPLPLCEDONGTIE---ESGLKQIRVDALLPKEQBGACSSIDGRVAV 220
QY 206 EYLDPSPE--VQKKYAFKCHRLKENNIEQIYPVNAISPHNIHTFATGSDGDFVNIWDP 263
DB 221 EYFDQGDYNSGKFAFCHRLKNTNLAYFVNSIEFSPRKKFLYTAGSDGIISCNWL 280
QY 264 FNKRLCOFHRY-PTSIASLAFSNDGTLTATASSYMEWDETHPEDG---IFIRQVTDATKPE 304
DB 281 QTRKKIKNFAKFNEDSVVVIACSDN--ILCLATS-----DDT 315

RESULT 9
S71241
probable RNA export protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 26-May-2000
C:Accession: S71241
R:Goodman, H.M.; Gallant, P.; Keifer-Higgins, S.; Rubenfield, M.; Church, G.M.
submitted to the EMBL Data Library, April 1996
A:Description: A 37.5 Kb sequence from Arabidopsis thaliana chromosome 1.
A:Reference number: S71240
A:Accession: S71241
A:Molecule type: DNA
A:Residues: 1-251 <GOO>
A:Cross-references: EMBL:U53501; NID:g1297184; PID:g1297188
C:Genetics:
A:Introns: 19/3; 48/3; 74/3; 139/2; 186/3; 213/3; 225/1
C:Superfamily: unassigned WD repeat proteins; WD repeat homology
F:68-101/Domain: WD repeat homology <WD1>

Query Match 19.9%; Score 352.5; DB 2; Length 251;
Best Local Similarity 30.8%; Pred. No. 1e-22;
Matches 86; Conservative 45; Mismatches 87; Indels 61; Gaps 6;
QY 4 SNEKLNQPPEDGISSVSKFSPNTSOFLLVSSWSDTSVRLYDVP-----ANSMLKYQHTG 57
DB 15 NKSEYVTPSPADSISSLSFSR-ADILVATSWDQVQWCEISRSGLASAPKASISHDQ 73
QY 58 AVLDCAFTDP-THAMSGGLDHLQKMHDLNTOENLVGTHDAPIRCVYCEVNVNMTGS 115
DB 74 PVLCSAMKDDGTTFVSGGCDKQAKWPLLSGGQPVTVAMHEGPTAAWAPIGKNLGTGS 133
QY 116 WDQTVKLMPTPCNAGTFQ---PEKVYTLVSGDRLIVGTAGRRVLDLNNMGVQORRE 175
DB 134 WDNTLKYMDTRQNPVHITQQLPKCYTLVSKVHPLMVVGTADRLNLIENLQNPQTEFKRIQ 193
QY 176 SSLKYQTCIRAFPNKQY--VLSIEGRVAVELDPSPVQKKYAFKCHRLKENNIEQ 233
DB 194 SPLKYQTCVTAAPPDQGGFLELACSLDAYIL----- 224
QY 234 IYPVNAISPHNIHTFATGSDGDFVNIWDPFNKRLCOF 272
DB 225 -----DGAFNFWDKSKQRLKVF 242

RESULT 10
T02818
polyA export protein PAXP [imported] - Leishmania major (strain Friedlin)
C:Species: Leishmania major
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 26-May-2000
C:Accession: F81458; T02818
R:Wylex, P.J.; Audleman, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magnus, C.;
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-
A:Reference number: A81455; MUID:99178987; PMID:10077609
A:Accession: F81458

R.Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena Res. 8, 205-213, 2001
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AE1810
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1227 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA074499.1; PID:g17135007; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr2800

Query Match 11.6%; Score 206; DB 2; Length 1227;
Best Local Similarity 24.8%; Pred. No. 1.9e-09;
Matches 84; Conservative 61; Mismatches 139; Indels 42; Gaps 15;
QY 15 DGISSVKPNTSQFLLVSSWMTSVRLYDV-PANSMRLKYQHTGAVLDCAF-YDPTAHWS 72
DB 769 DMVRCVAFSPD-GNTLASSAADHTIKLMDVSGKCLRTLKSGHTGWRVAFSADQOTLAS 827
QY 73 GGLDQLKHLNDLTDQ--ENLIVGTAGRRVLYVMDLRN-----MGYVQ 171
DB 887 IKTLGHTNEVCSVAFSPDQGLTACVSLDQSVRLNCRGTGCLKAWYGNTDWALPVAESP 946
QY 172 QRRESLKYQTCIRAFPNKOGYVLSIEGRVAVEY-IDPSPEVOKKYAF----- 221
DB 947 DRQLASSNDKTKLMDQWQTKYTSLEGTDFIYGFAPSDQSLASATDSVRLWN 1006
QY 222 ----KCHRLKNNIEQIYPVNAISPHNIHNTFATGSGDGFVNIWDPFNKRLCQPHRYPT 277
DB 1007 ISTGQCFCILLEHTDWWY---AVPHFPQCKIATGSDACTVKNLWISGTGCKLTLSEHSD 1063
QY 278 SIASLAFSNDGTTIAIASS 296
DB 1064 KILGNWAFSDGQLLASASA 1082

RESULT 14
AE1810
AD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AE1810
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena Res. 8, 205-213, 2001
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AE1810
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1227 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA077553.1; PID:g17135007; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr0029

Query Match 11.6%; Score 206; DB 2; Length 1227;
Best Local Similarity 24.8%; Pred. No. 3.3e-09;
Matches 84; Conservative 61; Mismatches 132; Indels 62; Gaps 15;
QY 17 ISSVKFSPNTSQFLLVSSWMTSVRLYDV-PANSMRLKYQHTGAVLDCAFY-DPTAHWSGGL 75
DB 903 VYSVAFSPD-SQLASGRDYITGLWNLKTECHPULRGHQIRSVAFHPDGGKILASGSA 961
QY 76 DQLKMDL-NTDQENLVGT---HDAPIRVCVEYCPVNVNVTGSDQTVKLMWDPRT-PCN 130

DB 962 DNTKLMDISITNHSKIRITLTGTNNWTVVFPSPDKHTLASSSEDRITRLWMDKDTGDCL 1021
QY 131 AGTFSPQPEKVYTLVSQDRLIV--GTAGRRVLWDLRNMGYVQORRESSLKYOTRCIRAF 188
DB 1022 QKLKGSHWVWTVAFSPDGRILASGSADSEIKIDVAS-----GKCLQTL 1066
QY 189 PNKQGYVLS---SIEGR-VAVEYLDPSPEVOKKYAFKCHLKNENIEQIYPVNAISPHN 244
DB 1067 TDQGMWVAFSLDGLTLLASASEDQTVKLMNLTGCVHTLKGE-KQVY---SVAFSP 1122
QY 245 IHNTFATGSGDGFVNIWDPFNKRLCQF-HRYPTSIASLAFSNDGTTIA----- 292
DB 1123 NGQIAASGSEDTTVKLMWISGSCVDTLKCHHTPAIRSVAFSPDGRLLASGEDEKTLQW 1182
QY 293 -----IASSYMYE-MDDTEHPEDGIFIROVTDAAE 320
DB 1183 DMQNCRLKTLKSPRYENMDITD-----ITGITDAE 1214

RESULT 15

H86292
F7H2.18 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Nov-2001
C:Accession: H86292
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizuar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: H86292
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-140 <STO>
A:Cross-references: GB:AE005172; NID:g8927663; PIDN:AAF82154.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 11.2%; Score 198.5; DB 2; Length 140;
Best Local Similarity 35.7%; Pred. No. 7.9e-10;
Matches 45; Conservative 26; Mismatches 46; Indels 9; Gaps 4;
QY 4 SNEFKLNQPPEDGISSVKFSPNTSQFLLVSSWMTSVRLYDVP-----ANSMLKYQHTG 57
DB 16 NNSYEITPNTDISISLSFSFK-ADILVATSDQVRCWEITRSDGISASEPKVNSHDQ 74
QY 58 AVLDCAFYD-THAWSGLDQLKMDH-LNTQENLVGTADAPIRVCVEYCPVNVNVTGS 115
DB 75 PVLCSAWKDDGTTVFTGGCDKQAKMWPILLSGAQSPSTVAMHDAFNFQIAPNGNLLVTGS 134
QY 116 WDQTVK 121
DB 135 WDKTLK 140

Search completed: November 12, 2003, 20:10:46
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 20:05:12 ; Search time 43 Seconds
(without alignments)
1210.752 Million cell updates/sec

Title: US-10-084-700-2
Perfect score: 1772
Sequence: 1 MTGSEFKLNQPPEDGSSV.....DGIFRQVDAETKPSCT 328

Scoring table: BLOSUM62
Gapop 10.0, Capext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

ALIGNMENTS

RESULT 1
AA06288
ID AAY06288 standard; Protein; 328 AA.

AC AAY06288;

DT 23-AUG-1999 (first entry)

DE Human BUB3, a substrate for BUB1A and BUB1B.

XX BUB3; human; protein kinase; checkpoint control; mitosis;
kinetochore; cell proliferation; cancer; tumour; therapy.

OS Homo sapiens.

XX WO9928334-A1.

XX 10-JUN-1999.

XX 01-DEC-1998; 98WO-US25415.

XX 01-DEC-1997; 97US-0067093.

XX (FOXC-) FOX CHASE CANCER CENT.

XX Chan G, Jablonski S, Yen T;

XX WPI; 1999-371094/31.

XX N-PSDE; AAY58972.

XX Human BUB genes and proteins involved in mitotic checkpoint control

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1772	100.0	328	20	Human BUB3, a subs
2	1772	100.0	328	23	Human BUB3 protein
3	1772	100.0	328	24	Human BUB3 protein
4	1772	100.0	332	21	Human cancer assoc
5	1763	99.5	328	20	Human huBUB3 prote
6	1050.5	59.3	328	22	Drosophila melanog
7	939	53.0	340	21	Arabidopsis thalia
8	705.5	39.8	166	23	Human ovarian anti
9	581.5	32.8	343	21	Zea mays protein f

Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Human 5' EST seque
Drosophila melanog
Human transport-as
Lung cancer associ
Zea mays protein f
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Zea mays protein f
Arabidopsis thalia
Novel human diagno
Human BUB3 protein
Drosophila melanog
Human 5' EST seque
Secreted protein 5
Arabidopsis thalia
Novel human diagno
Human colon cancer
Drosophila melanog
Human ORFX ORF3077
Human polypeptide,
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Human GTPase-assoc
Human GAP-1 prote
Human GTPase assoc
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia

PS Claim 32; Page 94-95; 99pp; English.

XX This sequence represents human BUB3, a protein that complexes with
 CC BUB1A (see AAY06286) and BUB1B (see AAY06287), novel kinases that play
 CC a pivotal role in mitotic checkpoint control. BUB3 is a substrate
 CC of these kinases. The BUB3 protein comprises 5 WD-40 motif repeats.
 CC It localises to the kinetochores during mitosis. BUB3 cDNA (see
 CC AAX58970) was isolated from a human EST database, with additional
 CC sequences obtained by 5'RACE. BUB genes (see AAX58970-72) and their
 CC encoded proteins (see AAY06286-88) disclosed in the invention provide
 CC valuable therapeutic targets for the design of antiproliferative
 CC agents which inhibit the aberrant cellular proliferation observed
 CC in tumour cells. Antibodies and oligonucleotide probes based on
 CC BUB are also useful in screening for therapeutics, diagnosis and
 CC identification of homologues.

XX SQ Sequence 328 AA;

Query Match 100.0%; Score 1772; DB 20; Length 328;
 Best Local Similarity 100.0%; Pred. No. 1.1e-177;
 Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTGSNEFKLNQPPEDGISSVKFSPNTSQFLVSSMDTSVRLYDVPANSMRLKYQHTGAVL 60
 Db 1 MTGSNEFKLNQPPEDGISSVKFSPNTSQFLVSSMDTSVRLYDVPANSMRLKYQHTGAVL 60
 Qy 61 DCAFYDPTAWSGGLDHLQKMDLNTDQENLVGTHDAPRCVCEYCEVNVMTGSDQTV 120
 Db 61 DCAFYDPTAWSGGLDHLQKMDLNTDQENLVGTHDAPRCVCEYCEVNVMTGSDQTV 120
 Qy 121 KLMDPRTPCNAGTFSQPEKVTYLSVSGDRLIVGTAGRRVLVMDLRNMGVYVQQRRESSLKY 180
 Db 121 KLMDPRTPCNAGTFSQPEKVTYLSVSGDRLIVGTAGRRVLVMDLRNMGVYVQQRRESSLKY 180
 Qy 181 QTRCIRAFPNKQGVLLSSIEGRVAVELDPSPEVQKKYAFKCHRLKENNIQIYFVNAI 240
 Db 181 QTRCIRAFPNKQGVLLSSIEGRVAVELDPSPEVQKKYAFKCHRLKENNIQIYFVNAI 240
 Qy 241 SFNHNHTFATGSGDGFVNIMDPPFNKKRLCQFHRYPTSIASLAFSNDGTTTIAISSYMYE 300
 Db 241 SFNHNHTFATGSGDGFVNIMDPPFNKKRLCQFHRYPTSIASLAFSNDGTTTIAISSYMYE 300
 Qy 301 MDDTEHPEDGIFIRQVTDATETPKSPCT 328
 Db 301 MDDTEHPEDGIFIRQVTDATETPKSPCT 328

RESULT 2

AAE25907
 ID AAE25907 standard; Protein; 328 AA.
 XX AAE25907;
 AC AAE25907;
 XX 15-NOV-2002 (first entry)
 DT 15-NOV-2002 (first entry)
 DE Human BUB3 protein.
 XX Human; cancer associated protein kinase; BUB3; restorative therapy;
 XX tumour; gene therapy; cancer; cytostatic; enzyme; chromosome 10.
 XX Homo sapiens.
 XX OS
 XX US6410312-B1.
 XX 25-JUN-2002.
 XX 10-DEC-1998; 98US-0208743.
 XX 19-DEC-1997; 97US-068102P.
 XX 30-DEC-1997; 97US-070182P.
 XX 11-JUN-1998; 98US-088991P.
 XX (CHIR) CHIRON CORP.

XX Seeley TW;

XX WPI: 2002-597913/64.
 XX N-PSDB; AAD42653.

XX Polynucleotides encoding a cancer associated human protein kinase BUB3
 PT (hUBUB3) protein, useful for the prevention, diagnosis and treatment of
 PT cancers -

XX Claim 1; Column 49-52; 47pp; English.

XX The invention relates to cancer associated human protein kinase BUB3
 CC (hUBUB3) protein and its nucleic acid sequence. The invention also
 CC relates to a method for assessing mutations and/or loss of the hUBUB3
 CC gene in human tumours. hUBUB3 DNA is used in the prevention, diagnosis
 CC and treatment of cancers associated with its inappropriate expression.
 CC hUBUB3 DNA and its complements may also be used as DNA probes and
 CC primers in diagnostic assays to detect and quantitate the presence of
 CC similar nucleic acids in samples, and therefore which patients may be
 CC in need of restorative therapy. The hUBUB3 polypeptides may also be
 CC used as antigens in the production of antibodies against hUBUB3 and in
 CC assays to identify modulators of hUBUB3 expression and activity. The
 CC anti-hUBUB3 antibodies and antagonists may also be used to down
 CC regulate expression and activity. The anti-hUBUB3 antibodies may also
 CC be used as diagnostic agents for detecting the presence of hUBUB3
 CC polypeptides in samples (e.g. by enzyme linked immunosorbent assay
 CC (ELISA)). hUBUB3 DNA is used in gene therapy. The present sequence is
 CC human BUB3 protein. hUBUB3 gene is located at chromosome 10.

XX SQ Sequence 328 AA;

Query Match 100.0%; Score 1772; DB 23; Length 328;
 Best Local Similarity 100.0%; Pred. No. 1.1e-172;
 Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTGSNEFKLNQPPEDGISSVKFSPNTSQFLVSSMDTSVRLYDVPANSMRLKYQHTGAVL 60
 Db 1 MTGSNEFKLNQPPEDGISSVKFSPNTSQFLVSSMDTSVRLYDVPANSMRLKYQHTGAVL 60
 Qy 61 DCAFYDPTAWSGGLDHLQKMDLNTDQENLVGTHDAPRCVCEYCEVNVMTGSDQTV 120
 Db 61 DCAFYDPTAWSGGLDHLQKMDLNTDQENLVGTHDAPRCVCEYCEVNVMTGSDQTV 120
 Qy 121 KLMDPRTPCNAGTFSQPEKVTYLSVSGDRLIVGTAGRRVLVMDLRNMGVYVQQRRESSLKY 180
 Db 121 KLMDPRTPCNAGTFSQPEKVTYLSVSGDRLIVGTAGRRVLVMDLRNMGVYVQQRRESSLKY 180
 Qy 181 QTRCIRAFPNKQGVLLSSIEGRVAVELDPSPEVQKKYAFKCHRLKENNIQIYFVNAI 240
 Db 181 QTRCIRAFPNKQGVLLSSIEGRVAVELDPSPEVQKKYAFKCHRLKENNIQIYFVNAI 240
 Qy 241 SFNHNHTFATGSGDGFVNIMDPPFNKKRLCQFHRYPTSIASLAFSNDGTTTIAISSYMYE 300
 Db 241 SFNHNHTFATGSGDGFVNIMDPPFNKKRLCQFHRYPTSIASLAFSNDGTTTIAISSYMYE 300
 Qy 301 MDDTEHPEDGIFIRQVTDATETPKSPCT 328
 Db 301 MDDTEHPEDGIFIRQVTDATETPKSPCT 328

RESULT 3

ABG75607
 ID ABG75607 standard; Protein; 328 AA.
 XX ABG75607;
 AC ABG75607;
 XX 15-APR-2003 (first entry)
 DT 15-APR-2003 (first entry)
 DE Human BUB3 protein (hUBUB3).
 XX Human; enzyme; BUB3; hUBUB3; cell cycle control; apoptosis;
 XX neoplastic tissue; neoplasia; kinase; cytostatic; kinase inhibitor;
 KW

cancer.
 Homo sapiens.
 US2002160403-A1.
 31-OCT-2002.
 27-FEB-2002; 2002US-0084700.
 11-JUN-1997; 97US-049068P.
 19-DEC-1997; 97US-068102P.
 30-DEC-1997; 97US-070182P.
 11-JUN-1998; 98US-088991P.
 10-DEC-1998; 98US-0208743.
 (CHIR) CHIRON CORP.
 Seeley TW;
 WPI; 2003-209220/20.
 N-PSDB; ABX15878.
 New isolated and purified hUBUB3 protein, useful for the diagnosis and treatment of disorders with mutation of the hUBUB3 protein, such as cancer of the lung, breast, colorectal, bladder, prostate, liver or stomach -
 Claim 2; Fig 2; 52pp; English.
 This invention relates to an isolated and purified hUBUB3 protein which is involved in cell cycle control and apoptosis. The invention also comprises a method for diagnosing a neoplastic tissue of a human, and a method for aiding in the diagnosis or prognosis of neoplasias in a human. Also disclosed are methods for identifying compounds which decrease the kinase activity of a hUBUB1-hUBUB3 complex. The hUBUB protein of the invention may have cytoskeletal activity and the DNA sequence encoding it may be used in antisense-therapy as a kinase inhibitor. The methods and compositions are useful for the diagnosis and treatment of disorders with mutation of the hUBUB3 protein, such as cancer of the lung, breast, colorectal, bladder, prostate, liver or stomach. The present sequence represents the human hUBUB3 protein used in the methods of the invention.
 Query Match 100.0%; Score 1772; DB 24; Length 328;
 Best Local Similarity 100.0%; Pred. No. 1.1e-177;
 Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTGSEFLKQPPEDGSSVVFSPNTSOFLLVSSMDTSVRLVDVPSNMLKXQHTGAVL 60
 DB 1 MTGSEFLKQPPEDGSSVVFSPNTSOFLLVSSMDTSVRLVDVPSNMLKXQHTGAVL 60
 QY 61 DCAFYDTHAMSGGLDHLKMDLNTDQENLVGTHDAPRCVCEYCEPVNVMVGTGSDQTV 120
 DB 61 DCAFYDTHAMSGGLDHLKMDLNTDQENLVGTHDAPRCVCEYCEPVNVMVGTGSDQTV 120
 QY 121 KLMDPRTPCNAGTSPQEKVYVTLVSQDRLVGTAGRVLVMDLRNMGVYVQRRSSLYK 180
 DB 121 KLMDPRTPCNAGTSPQEKVYVTLVSQDRLVGTAGRVLVMDLRNMGVYVQRRSSLYK 180
 QY 181 QTRCIRAPFNKQGVYVSSIEGRVAVVEYLDPSPEVQKKYAFKCHRLKNNIEQIYVNVAI 240
 DB 181 QTRCIRAPFNKQGVYVSSIEGRVAVVEYLDPSPEVQKKYAFKCHRLKNNIEQIYVNVAI 240
 QY 241 SFNHNHTATGSGDGFVNTWDPNKKLCQFHYPTSTASAFSNDGTTLAIASSYMYE 300
 DB 241 SFNHNHTATGSGDGFVNTWDPNKKLCQFHYPTSTASAFSNDGTTLAIASSYMYE 300
 QY 301 MDDTEHPEDGIFIRQVTDATKPKSPCT 328
 DB 301 MDDTEHPEDGIFIRQVTDATKPKSPCT 328

RESULT 4
 AAB43856
 ID AAB43856 standard; Protein; 332 AA.
 XX
 AC AAB43856;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated protein sequence SEQ ID NO:1301.
 XX
 DE Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnary; immunomodulator;
 KW anti-diabetic; anti-asthmatic; anti-rheumatic; anti-arthritic; antiviral;
 KW anti-inflammatory; antithyroid; antiallergic; antibacterial; cardiac;
 KW dermatological; neuroprotective; thrombolytic; coagulant; neoplastic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disease; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.
 XX
 OS Homo sapiens.
 XX
 FN WO200055350-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05882.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2000-587533/55.
 DR N-PSDB; AAC78065.
 XX
 PT Novel isolated nucleic acids comprising sequences encoding peptides
 XX useful for treating or diagnosing e.g. cancer -
 XX
 PS Claim 11; Page 1946-1947; 2352pp; English.
 XX
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnary; immunomodulator;
 CC anti-diabetic; anti-asthmatic; anti-rheumatic; anti-arthritic;
 CC anti-inflammatory; neuroprotective; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
 CC neoplastic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.
 XX
 XX Sequence 332 AA;

Query Match 100.0%; Score 1772; DB 21; Length 332;
 Best Local Similarity 100.0%; Pred. No. 1.2e-177;
 Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGNEFKLNQPPEDGSSVVKFSFNTSQFLVSSWDTSVRLYDVPANSMRLKYQHTGAVL 60
 DB 5 MTGNEFKLNQPPEDGSSVVKFSFNTSQFLVSSWDTSVRLYDVPANSMRLKYQHTGAVL 64
 QY 61 DCAFYDPTAHWSGGDLHQLKXHDLNTDQENLVGTHDAPIRCVCEYCEPVNVMVTGSDQTV 120
 DB 65 DCAFYDPTAHWSGGDLHQLKXHDLNTDQENLVGTHDAPIRCVCEYCEPVNVMVTGSDQTV 124
 QY 121 KLMDPRTPCNAGTFSQPEKVTYLSVSGDRLIVGTAGRRVLVMDLRNMGYYVQQRRESSLKY 180
 DB 125 KLMDPRTPCNAGTFSQPEKVTYLSVSGDRLIVGTAGRRVLVMDLRNMGYYVQQRRESSLKY 184
 QY 181 QTRCIRAFPNKQGYVLSIEGRVAVELYDPSPEVQKKYAFKCHRLKNNIEQIYPVNAI 240
 DB 185 QTRCIRAFPNKQGYVLSIEGRVAVELYDPSPEVQKKYAFKCHRLKNNIEQIYPVNAI 244
 QY 241 SFHNIHTFATGSGDGFVNIWDPFNKRLCOFHRYPTSIASLAFSNDGTTLAIASSYMYE 300
 DB 245 SFHNIHTFATGSGDGFVNIWDPFNKRLCOFHRYPTSIASLAFSNDGTTLAIASSYMYE 304
 QY 301 MDDTEHPEDGIFIRQVTDATKPKSPCT 328
 DB 305 MDDTEHPEDGIFIRQVTDATKPKSPCT 332

RESULT 5
 ID AAW90122 standard; Protein; 328 AA.

AC AAW90122;
 DT 12-APR-1999 (first entry)
 DE Human huBUB3 protein.
 KW BUB1; huBUB3; human; huBUB1; cell cycle; apoptosis; tumour;
 KW cancer; neoplasia; psoriasis; diagnosis; therapy.
 OS Homo sapiens.
 XX WO9856910-A1.
 XX 17-DEC-1998.
 XX 11-JUN-1998; 98WO-US11882.
 XX 30-DEC-1997; 97US-0070182.
 XX 11-JUN-1997; 97US-0049068.
 XX 19-DEC-1997; 97US-0068102.
 XX (CHIR) CHIRON CORP.
 XX Seeley T;
 XX WPI; 1999-070322/06.
 XX N-PSDB; AAV74173.
 XX New human BUB1 protein - useful for diagnosis and treatment of
 XX cancer
 XX Disclosure; Page 56; 68pp; English.
 XX This is the amino acid sequence of human huBUB3 protein. A new
 XX human BUB1 protein, designated huBUB1 (see AAW90122), binds to huBUB3.
 XX huBUB1 and the huBUB1-huBUB3 complex have kinase activity. The
 XX huBUB1 gene (see AAV74172) is mutated in cancer cells. Agents that
 XX interfere with huBUB1-huBUB3 interaction, or which reduce huBUB1
 XX kinase activity are used to increase the sensitivity of tumours to
 XX anticancer agents, or to treat other hyperproliferative diseases
 XX such as psoriasis.
 XX Sequence 328 AA;

Query Match 99.5%; Score 1763; DB 20; Length 328;
 Best Local Similarity 99.7%; Pred. No. 1e-176;
 Matches 327; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MTGNEFKLNQPPEDGSSVVKFSFNTSQFLVSSWDTSVRLYDVPANSMRLKYQHTGAVL 60
 DB 1 MTGNEFKLNQPPEDGSSVVKFSFNTSQFLVSSWDTSVRLYDVPANSMRLKYQHTGAVL 60
 QY 61 DCAFYDPTAHWSGGDLHQLKXHDLNTDQENLVGTHDAPIRCVCEYCEPVNVMVTGSDQTV 120
 DB 61 DCAFYDPTAHWSGGDLHQLKXHDLNTDQENLVGTHDAPIRCVCEYCEPVNVMVTGSDQTV 120
 QY 121 KLMDPRTPCNAGTFSQPEKVTYLSVSGDRLIVGTAGRRVLVMDLRNMGYYVQQRRESSLKY 180
 DB 121 KLMDPRTPCNAGTFSQPEKVTYLSVSGDRLIVGTAGRRVLVMDLRNMGYYVQQRRESSLKY 180
 QY 181 QTRCIRAFPNKQGYVLSIEGRVAVELYDPSPEVQKKYAFKCHRLKNNIEQIYPVNAI 240
 DB 181 QTRCIRAFPNKQGYVLSIEGRVAVELYDPSPEVQKKYAFKCHRLKNNIEQIYPVNAI 240
 QY 241 SFHNIHTFATGSGDGFVNIWDPFNKRLCOFHRYPTSIASLAFSNDGTTLAIASSYMYE 300
 DB 241 SFHNIHTFATGSGDGFVNIWDPFNKRLCOFHRYPTSIASLAFSNDGTTLAIASSYMYE 300
 QY 301 MDDTEHPEDGIFIRQVTDATKPKSPCT 328
 DB 301 MDDTEHPEDGIFIRQVTDATKPKSPCT 328

RESULT 6
 ID ABB59183 standard; Protein; 326 AA.

AC ABB59183;
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 4341.
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 OS Drosophila melanogaster.
 XX WO200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US09231.
 XX 23-MAR-2000; 2000US-191637P.
 XX 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 XX N-PSDB; ABL03286.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 XX genes from Drosophila and for elucidating cell signalling and cell-cell
 XX interactions -
 XX Disclosure; SEQ ID NO 4341; 2ipp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 XX capable of detecting 1000 or more genes from Drosophila. The invention is
 XX useful in developmental biology and in elucidating cell signalling and
 XX cell-cell interactions in higher eukaryotes for the development of
 XX insecticides, therapeutics and pharmaceutical drugs. The invention
 XX discloses genomic DNA sequences (ABLU16176-ABLU30511), expressed DNA
 XX sequences (ABLU1840-ABLU16175) and the encoded proteins

```

CC (ABBS7737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 326 AA;

Query Match          59.3%; Score 1050.5; DB 22; Length 326;
Best Local Similarity 60.6%; Pred. No. 1.1e-101;
Matches 195; Conservative 49; Mismatches 75; Indels 3; Gaps 2;

Qy 6 EFKLNQPEEGISVSKFSPNTSOFLLVSSWDTSVRLYDVPANSMRLKYQHTGAVLDCAFY 65
Dy 5 EFKLNQPEELISAVKFGKSNQYMAASSWDGILRFDYDVPANQLRQKFPQDAPLLDCAPM 64
Qy 66 DPHWAGGLDHLQKMDLNTDOENLVGTHDAPIRCEVCPEVNVNMTGSDQTVKLMDP 125
Dy 65 DIVHVGSLDNQLRFLPDVNTQAESIIIGAHEEPIRCVHAHYVNGILTGMONTVKLMDM 124
Qy 126 RTPNAGTFSQPR-KVYTLVSGDRLIVGTAGRVLVMDLRNMGYVQQRRESSLKYQTEC 184
Dy 125 REKFCVGTFFQNNQKVSMSVIDEKIVVATSDRKLVDLWDRKXDSYIMKRESSLKYQTRC 184
Qy 185 IRAPPNKGQVLSIEGRVAVEYLDPSPEVQKKYAFKCHRLKNNIEQIYPVNAISFHN 244
Dy 185 IRLFPNKEGYVMSIEGRVAVEYLDHDPVEVQRKFAFKCHRNREQNEIQIYPVNAISFHN 244
Qy 245 IHNTFATGGGDFVNDPNNKPLCOFHYPTSIASLAFSNDGTTLAASSVYVEMDDT 304
Dy 245 VYQFATGGSDGIWNIWDGNNKRLCQFHEYDTSITLNFSDGSALAGCSYLDQLPET 304
Qy 305 EH--PEDGIFIRQVTDATKPK 324
Dy 305 PATVPHPAIYIRYPTDQETQK 326

RESULT 7
AAG16894
ID AAG16894 standard; Protein; 340 AA.
XX
AC AAG16894;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 17711.
XX
KW protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0132864.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139432.
PR 17-JUN-1999; 99US-0139432.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.

```

```

PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148585.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151055.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156536.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

```

```

PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 53.0%; Score 939; DB 21; Length 340;
Best Local Similarity 54.1%; Pred. No. 6.2e-90;
Matches 178; Conservative 54; Mismatches 91; Indels 6; Gaps 2;

QY 1 MTGSEFKLNQPPEDGIGSSVKSPNTSQFLVSSMDTSVRLYDVDPANSMRLKYQHTGAVL 60
DB 4 VTPSAGRELSNPPSDGINSRLRFS-NNSDHLVSSMDKRVRLYDVSTNSLKGFLHGGAVL 62
QY 61 DCAFYDPTHAMSGGLDHLQKMHDLNTDQENLVGTHDAPIRCVVEYCEPVNVMVTGSDQTV 120
DB 63 DCCFHDDPSGFSVADYKVRIRIVFNVGKEDILGTHDKAVRCVEYSYAAQGVITGSMDKTV 122
QY 121 KLWDPR-----TFCNAGTFSQPEKYVTLVSVDGLTGTAGRVLYMDLRNMGYVOORSE 175
DB 123 KWDPRGASGPRTQGTILQPERVYSLSVGLVHRLVATAGRVNIDLRNMSQPQRRE 182
QY 176 SSLKYQTCIRAFPNKQGYLVSSIEGRVAVELYDPSPEVQKKYAFKCHRLKENNIEQIV 235
DB 183 SSLKYQTCRCVCPNGTGYALSSVEGRVAMEFOLSEAAQKKAFFKCHKSEAGRDIVY 242
QY 236 PVNAISFHNHTFATGSDGFVNIWDPKRLQLQFHYPTSIASLAFNSDGTTLAIAS 295
DB 243 PVNSIAFHPIYGTFTATGGDGFVNIWGNKRLYQYSKYPTISALSFSRDQQLLAVAS 302
QY 296 SYMYEMDDTEHPEDGIFIRQVTDATPK 324
DB 303 SYTFEGEKSQPEPAIFVRSVNEIEVKPK 331

RESULT 8
ABP43220
ID ABP43220 standard; Protein; 166 AA.
XX
AC ABP43220;
XX
DT 22-AUG-2002 (first entry)
XX
DE Human ovarian antigen HVVDX90, SEQ ID NO:4352.
XX
KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive.
OS Homo sapiens.
XX
XX WO200200677-A1.
PN

```


XX PD 03-JAN-2002.
 XX PF 07-JUN-2001; 2001MO-US18569.
 XX PR 07-JUN-2000; 2000US-209467P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Birse CE, Rosen CA;
 XX DR WPI: 2002-147878/19.
 XX DR N-PSDB; ABQ56297.
 XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX Claim 11; SEQ ID No 4352; 2922pp; English.
 XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 98% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 166 AA;
 S0 Query Match 39.8%; Score 705.5; DB 23; Length 166;
 Best Local Similarity 93.7%; Pred. No. 8e-66;
 Matches 133; Conservative 1; Mismatches 3; Indels 5; Gaps 1;
 Qy 184 CIRAFNPKQGVLSSTEGRAVEVLDPSPVQKKYAFKCHRLKENNIEQIYFVNAISPH 243
 Db 29 CVRX-----GVLSSTEGRAVEVLDPSPVQKKYAFKCHRLKENNIEQIYFVNAISPH 83
 Qy 244 NIHNTATGSGDGFVNWDPFNKKRLCQPHRYPTSIASLAFSNDGTTLAIASSYMYEMDD 303
 Db 84 NIHNTATGSGDGFVNWDPFNKKRLCQPHRYPTSIASLAFSNDGTTLAIASSYMYEMDD 143
 Qy 304 TEHPEDGIFIRQVDAETKPS 325
 Db 144 TEHPEDGIFIRQVDAETKPS 165
 RESULT 9
 AAG33215
 ID AAG33215 standard; Protein; 343 AA.

XX AC AAG33215;
 XX DT 18-OCT-2000 (first entry)
 XX DE Zea mays protein fragment SEQ ID NO: 40212.
 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence; corn.
 XX OS Zea mays subsp. mays.
 XX PN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0301439.
 XX PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126795.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131149.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134257.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.

```

PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142802.
PR 09-JUL-1999; 99US-0142940.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 18-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.

PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0158295.
PR 14-OCT-1999; 99US-0158329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      32.8%; Score 581.5; DB 21; Length 343;
Best Local Similarity 38.8%; Pred. No. 2.7e-52;
Matches 128; Conservative 64; Mismatches 122; Indels 15; Gaps 8;

QY 4 SNEFKLNQPPEDGISSVFKSPNTSQFLLVSSDWTSLVLYDVNPANMRK--YQHTGAVLD 61
Db 14 NKSEILNFGDSVSSLSFSPK-SNLLVATSWDNQVRCWEIVGNSQPKASISHDQPVLC 72
QY 62 CAFYDP-THAWSGGLDHLKMD-LNTDOENLACTHDAPIRCVEYCEPVANVMTGSNDQT 119
Db 73 SAWKDDGTTVFSGGCDQVKNWPLLSGGQPTVAMHDAPVKEVAWIPQMLLVSGSDXT 132
QY 120 VKLMDPRTPCNAGTFSQPEKVTYLSVSGDLIVGTAGRRVLVMDLRNMGVVOQRRESSLK 179
Db 133 LRYWDTFQSNPVHIIQQLPERCYALTNYPLMIIVGTADRLVNVNQLNPQTEFRIGSPK 192
QY 180 YQTRCTRAFENKQGVLSISIEGRVAVVEYLDPSPEVQKKYAFKRLKENNIEQIYVNA 239
Db 193 YQTRCLAAFPDQGGFLVSGIEGRVGVHHDDAQ--QSKNFTFKCHR--EGN--DIFSUNS 246

```


PR	10-AUG-1999;	99US-0148171.	DB	1	MSLVGHLVATAGRHVNIYDLNMSQEQRESSLKYQTRCVRCPNGTGYALSVEGR	60
PR	11-AUG-1999;	99US-0148319.	QY	203	VAVEYLDPSPEVOKKYAPKCHRLKENNIEBOLYPVNAISPHNIHTFATGSGGFVNIWD	262
PR	12-AUG-1999;	99US-0148341.	DB	61	VAMEFFDLSEAAQAKYAFKCHRKSEAGRDIVPVNSIAFHPIYGTGATGGCGGFVNIWD	120
PR	13-AUG-1999;	99US-0148585.	QY	263	PFNKRLCQFHRYPSTIASAFNDGTTTIAISSYMYEMDTHPHDGPPIROVTDATK	322
PR	16-AUG-1999;	99US-0149368.	DB	121	GNNKKRLYQYKYPTSISSALSFSDGQLLAVASSYTFEGEKSEPEAFVRSVNEIEVK	180
PR	17-AUG-1999;	99US-0149175.	QY	323	PK 324	
PR	18-AUG-1999;	99US-0149722.	DB	181	PK 182	
PR	20-AUG-1999;	99US-0149723.				
PR	20-AUG-1999;	99US-0149929.				
PR	23-AUG-1999;	99US-0149902.				
PR	23-AUG-1999;	99US-0149930.				
PR	25-AUG-1999;	99US-0150566.				
PR	26-AUG-1999;	99US-0150884.				
PR	27-AUG-1999;	99US-0151065.				
PR	27-AUG-1999;	99US-0151066.				
PR	27-AUG-1999;	99US-0151080.				
PR	30-AUG-1999;	99US-0151080.				
PR	31-AUG-1999;	99US-0151438.				
PR	01-SEP-1999;	99US-0151930.				
PR	07-SEP-1999;	99US-0152363.				
PR	10-SEP-1999;	99US-0153070.				
PR	13-SEP-1999;	99US-0153758.				
PR	15-SEP-1999;	99US-0154018.				
PR	16-SEP-1999;	99US-0154039.				
PR	20-SEP-1999;	99US-0154779.				
PR	22-SEP-1999;	99US-0155139.				
PR	23-SEP-1999;	99US-0155486.				
PR	24-SEP-1999;	99US-0155659.				
PR	28-SEP-1999;	99US-0156458.				
PR	29-SEP-1999;	99US-0156596.				
PR	04-OCT-1999;	99US-0157117.				
PR	05-OCT-1999;	99US-0157753.				
PR	06-OCT-1999;	99US-0157865.				
PR	07-OCT-1999;	99US-0158029.				
PR	08-OCT-1999;	99US-0158232.				
PR	12-OCT-1999;	99US-0158369.				
PR	13-OCT-1999;	99US-0159293.				
PR	13-OCT-1999;	99US-0159294.				
PR	13-OCT-1999;	99US-0159295.				
PR	14-OCT-1999;	99US-0159329.				
PR	14-OCT-1999;	99US-0159330.				
PR	14-OCT-1999;	99US-0159331.				
PR	14-OCT-1999;	99US-0159637.				
PR	14-OCT-1999;	99US-0159638.				
PR	18-OCT-1999;	99US-0159584.				
PR	18-OCT-1999;	99US-0160741.				
PR	21-OCT-1999;	99US-0160767.				
PR	21-OCT-1999;	99US-0160768.				
PR	21-OCT-1999;	99US-0160770.				
PR	21-OCT-1999;	99US-0160814.				
PR	21-OCT-1999;	99US-0160815.				
PR	22-OCT-1999;	99US-0160980.				
PR	22-OCT-1999;	99US-0160981.				
PR	22-OCT-1999;	99US-0160989.				
PR	25-OCT-1999;	99US-0161404.				
PR	25-OCT-1999;	99US-0161405.				
PR	25-OCT-1999;	99US-0161406.				
PR	26-OCT-1999;	99US-0161359.				
PR	26-OCT-1999;	99US-0161360.				
PR	26-OCT-1999;	99US-0161361.				
PR	28-OCT-1999;	99US-0161920.				
PR	28-OCT-1999;	99US-0161992.				
PR	28-OCT-1999;	99US-0161993.				
PR	29-OCT-1999;	99US-0162142.				
QY	143	LSVSGDRLIVGTAGRHVNIYDLNMSQEQRESSLKYQTRCVRCPNGTGYALSVEGR	202			

Query Match 32.7%; Score 579; DB 21; Length 191;
 Best Local Similarity 59.3%; Pred. No. 2, 1e-52;
 Matches 108; Conservative 27; Mismatches 47; Indels 0; Gaps 0;
 143 LSVSGDRLIVGTAGRHVNIYDLNMSQEQRESSLKYQTRCVRCPNGTGYALSVEGR 202

PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 24-JUN-1999; 99US-0140823.
 PR 28-JUN-1999; 99US-0140891.
 PR 29-JUN-1999; 99US-0140891.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142820.
 PR 12-JUL-1999; 99US-0142877.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 18-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144352.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145092.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145513.
 PR 27-JUL-1999; 99US-0145518.
 PR 27-JUL-1999; 99US-0145519.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146389.
 PR 02-AUG-1999; 99US-0146389.
 PR 04-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 25-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 27-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match

32.6%; Score 578.5; DB 21; Length 349;

PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 08-AUG-1999; 99US-0147493.
 PR 08-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149829.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150366.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 32.6%; Score 578.5; DB 21; Length 349;
 Best Local Similarity 38.7%; Pred. No. 5.9e-52;
 Matches 129; Conservative 60; Mismatches 125; Indels 19; Gaps 8;

Qy 4 SNEFKLNQPPEDGISSVKFSPNTSQELVSSWMTSVRLYDVP-----ANSRLKYQHTG 57
 Db 15 NKSYEVTPSPADSISSLSFSPR-ADILVATSWDNQVRCWEISRSGASLASAPKASISHQ 73
 Qy 58 AVLDCAFYDP-THAWGGDLHQLKMD-LNTDQENLVGTHDAPIRCVEYCPEVNVMTGS 115
 Db 74 PVLCSAWKDDGTVFSGGCKQAKMPLLSGGOPVTVMHEGPIAAKMANIPGNLLATGS 133
 Qy 116 WDQTVKLMDPRTFCNAGTFSQPEKYVTLVSGDRLLVGTAGRRVLVWDLRNMGYVQORRE 175
 Db 134 WDKTLKYWDTRQONPVHTQOLPKCYTSLSVKHPVVGADRNLI VFNLQNPOTEFKRIQ 193
 Qy 176 SSLKYOTRCIRAPENKOGYVLSIEGRVAVEVLDPSPEVQKKYAPKCHRLKENNTEIY 235
 Db 194 SPLKYQTRCVTAPFDQGGPLVGSIEGRVGVHLLDSQ--QSKNFTFCHR--DGN--DIY 247
 Qy 236 PVNAISPHNIHTFATGSDGFVNIWDPFNKRLCOFHRYPTSIASIAFSNDGTTLIAAS 295
 Db 248 SVNSLNFHPVHGTFATAGSDGAFNWDKSKQRLKAMRCNQIPCCSFNHDGSIYVAA 307
 Qy 296 SYNYEMDDTEH-----PEDGIFIRQVDAETPK 324
 Db 308 CYDWSKGAENHPATAKSSIFLHPQSEVYKAK 340

RESULT 13
 ID AAY12333 standard; Protein; 107 AA.
 XX AC AAY12333;
 XX DT 17-JUN-1999 (first entry)
 XX DE Human 5' EST secreted protein SEQ ID NO:364.
 KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition.
 XX OS Homo sapiens.
 XX PN WO9906548-A2.
 XX PD 11-FEB-1999.
 XX PF 31-JUL-1998; 98WO-1B01222.
 XX PR 01-AUG-1997; 97US-0905135.
 XX PA (GBST) GENSET.
 XX PI Duclert A, Dumas Milne Edwards J, Lacroix B;
 XX WPI; 1999-153778/13.
 DR

DR N-PSDB; AAX41166.
XX New nucleic acids encoding human secreted proteins - obtained from
PT CDNA libraries prepared from e.g. liver, ovary, brain, prostate,
PT kidney, lung, umbilical cord, placenta and colon tissue
XX
XX Claim 27; Page 697; 824pp; English.
XX
XX AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAY12261 to
CC AAY12514, respectively. The proteins given represent the signal peptide
CC and an N-terminal fragment of a secreted protein. The nucleic acid
CC sequences can be used for producing secreted human gene products. They
CC can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell.
XX
XX Sequence 107 AA;
SQ
Query Match 32.2%; Score 571; DB 20; Length 107;
Best Local Similarity 98.1%; Pred. No. 6e-52;
Matches 104; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MTGSEFKLNQPPEDGIGSSVKFSPNTSQFLVSSMDTSVRLYDVPANSMRLKYQHTGAVL 60
DB 1 MTGSEFKLNQPPEDGIGSSVKFSPNTSQFLVSSMDTSVRLYDVPANSMRLKYQHTGAVL 60
QY 61 DCAFYDPTHAWSGGLDHLKMDLNTDQENLVGTHDAPIRCVCEYCP 106
DB 61 DCAFYDPTHAWSGGLDHLKMDLNTDQENLVGTHDAPIRCVCEYCP 106
RESULT 14
ABB64713
ID ABB64713 standard; Protein; 346 AA.
XX
XX ABB64713;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 20931.
DE Drosophila melanogaster polypeptide; cell signalling; insecticide;
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PMD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX N-PSDB; ABL08816.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 20931; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AB857737-AB872072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 346 AA;
SQ
Query Match 29.4%; Score 521.5; DB 22; Length 346;
Best Local Similarity 36.4%; Pred. No. 5.8e-46;
Matches 123; Conservative 59; Mismatches 133; Indels 23; Gaps 11;
QY 5 NEFKLNQPPEDGIGSSVKFSPNTSQ--FLVSSMDTSVRLYDVPANSMRLKYQHT--GAV 59
DB 12 NDFEVASFPDDSVSALESPSTVQKNFLVAGSWDSTVRCWEVEQNGATVPKSMKTMGGPV 71
QY 60 LD-CAFYDPTHAWSGGLDHLKMDLNTDQENLVGTHDAPIRCVCEY--PEVNVMVTGSM 116
DB 72 LDVCSDDSGSKVFVASCDDQKVLMDLADQVQVAHDDGVPKTCMVKGPTTYTCLMTGSM 131
QY 117 DQTVKLMPTPCNAGTFSPQEKVYTLVSGDRLIVGTAGRVLMVMDLRNNGYVQQRRES 176
DB 132 DXTLKFMDTRSPNPMWTLNLPERCYCADVEYPMVAVGTANRGLIYSLSQNSPTVEYKQES 191
QY 177 SLKYOTRCIRAPFNKQ----GYVLSSTEGRAVEYLDPSPEVQKKYAFKCHRLK-ENNI 231
DB 192 PLKYQHRATISIFRDKKKEFTGALSGTEGRAVIAQVNPNGN--PKDNFTFKCHRTTGTSGY 249
QY 232 EGIYVNVNAISFHNHNTATGSGDFVNIWDPFNKRLCQFHYPTSIASLAFSNDGTTL 291
DB 250 QDIYAVNDIARFHVHGTTLTVSGDFTSFWDKARTKLKSETWDQISITKCGFNANGQIF 309
QY 292 ATASSYMYE----MDDTEHPEDGIFIRQVTDATKPK 324
DB 310 AVAVGYDWSKGHEYNPAKPKQ--IFLRSCYD-ELKPR 344
RESULT 15
AAY31647
ID AAY31647 standard; Protein; 368 AA.
XX
XX AAY31647;
XX
XX 02-NOV-1999 (first entry)
XX
XX Human transport-associated protein-9 (TRANP-9).
DE Human transport-associated protein-9 (TRANP-9).
XX
XX Transport-associated protein; TRANP; nuclear pore; nuclear transport;
XX vesicle trafficking; cancer; cystic fibrosis; multidrug resistance;
XX hypercholesterolaemia; diagnosis; treatment.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 101..115
XX FT /note= "Beta transducin family Trp-Asp repeat"
XX FT Region 144..158
XX FT /note= "Beta transducin family Trp-Asp repeat"
XX
XX WO9941373-A2.
XX
XX 19-AUG-1999.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 20:07:58 ; Search time 21 Seconds
(without alignments)
660.855 Million cell updates/sec

Title: US-10-084-700-2

Perfect score: 1772
Sequence: 1 MTGSNEFKLNQPPEDGISSV.....DGFIRQVTDARTKPKSPCT 328

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1772	100.0	328	4	US-09-095-881-4
2	258	14.6	59	4	US-09-247-155-157
3	197.5	11.1	251	4	US-09-291-170A-13
4	197.5	11.1	251	4	US-09-724-884-13
5	176.5	10.0	375	3	US-09-063-743-1
6	176.5	10.0	375	4	US-09-590-340-1
7	169.5	9.6	250	4	US-09-291-170A-12
8	169.5	9.6	250	4	US-09-724-884-12
9	169.5	9.6	340	1	US-08-190-802A-38
10	169.5	9.6	340	3	US-08-477-346-38
11	169.5	9.6	340	4	US-08-473-089-38
12	169.5	9.6	340	4	US-08-487-072A-38
13	169.5	9.6	340	4	US-09-245-039-1
14	169.5	9.6	704	1	US-08-188-582-18
15	169.5	9.6	704	1	US-08-646-715-18
16	166	9.4	318	1	US-08-190-802A-33
17	166	9.4	318	3	US-08-477-346-33
18	166	9.4	318	4	US-08-473-089-33
19	166	9.4	318	4	US-08-487-072A-33
20	163	9.2	713	1	US-08-190-802A-63
21	163	9.2	713	3	US-08-477-346-63
22	163	9.2	713	4	US-08-473-089-63
23	163	9.2	713	4	US-08-487-072A-63
24	163	9.2	713	4	US-09-108-857-3
25	160.5	9.1	340	1	US-08-190-802A-42
26	160.5	9.1	340	3	US-08-477-346-42
27	160.5	9.1	340	4	US-08-473-089-42

28	160.5	9.1	340	4	US-08-487-072A-42	Sequence 42, Appl
29	160.5	9.1	732	4	US-08-914-999-8	Sequence 8, Appl
30	158	8.9	2629	2	US-08-751-189-4	Sequence 4, Appl
31	158	8.9	2629	2	US-09-060-836-4	Sequence 4, Appl
32	158	8.9	2629	3	US-09-184-445-4	Sequence 4, Appl
33	156.5	8.8	326	1	US-08-190-802A-39	Sequence 39, Appl
34	156.5	8.8	326	3	US-08-477-346-39	Sequence 39, Appl
35	156.5	8.8	326	4	US-08-473-089-39	Sequence 39, Appl
36	156.5	8.8	326	4	US-08-487-072A-39	Sequence 39, Appl
37	155	8.7	798	1	US-08-190-802A-64	Sequence 64, Appl
38	155	8.7	798	1	US-08-308-818-2	Sequence 68, Appl
39	155	8.7	798	2	US-08-477-346-64	Sequence 2, Appl
40	155	8.7	798	3	US-08-473-089-64	Sequence 64, Appl
41	155	8.7	798	4	US-08-473-089-64	Sequence 64, Appl
42	155	8.7	798	4	US-08-473-089-68	Sequence 68, Appl
43	155	8.7	798	4	US-08-487-072A-64	Sequence 64, Appl
44	155	8.7	798	4	US-08-487-072A-68	Sequence 68, Appl
45	155	8.7	798	4	US-08-487-072A-68	Sequence 68, Appl

ALIGNMENTS

RESULT 1
US-09-095-881-4
; Sequence 4, Application US/09095881
; Patent No. 6489137
; GENERAL INFORMATION:
; APPLICANT: Seeley, Todd
; TITLE OF INVENTION: DETECTION OF LOSS OF THE WILD-TYPE
; FILE REFERENCE: 1405.003 / 200130.437
; CURRENT APPLICATION NUMBER: US/09/095.881
; CURRENT FILING DATE: 1998-06-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-095-881-4

Query Match 100.0%; Score 1772; DB 4; Length 328;
Best Local Similarity 100.0%; Pred. No. 1.2e-186; Indels 0; Gaps 0;
Matches 328; Conservative 0; Mismatches 0

Qy	1	MTGSNEFKLNQPPEDGISSVKFSFNTSQFLVSSWDTSVRLYDVPANSMRLKYQHTGAVL	60
Db	1	MTGSNEFKLNQPPEDGISSVKFSFNTSQFLVSSWDTSVRLYDVPANSMRLKYQHTGAVL	60
Qy	61	DCATYDPTHAMSGGLDHLKXMDLNTDOENLVGTHDAPICVEXCEPVNVMVVGSDQTV	120
Db	61	DCATYDPTHAMSGGLDHLKXMDLNTDOENLVGTHDAPICVEXCEPVNVMVVGSDQTV	120
Qy	121	KLMDPRTPCNAGTFSQPEKVTYLLSVSGDRLLVGTAGRRVLVWDLRNNGYVQQRRESSLKY	180
Db	121	KLMDPRTPCNAGTFSQPEKVTYLLSVSGDRLLVGTAGRRVLVWDLRNNGYVQQRRESSLKY	180
Qy	181	QTRCIRAPPNKGQVVLSSIEGRVAVELDPSPEVQKKYAFKCHRLKENNIEQIYPNVAI	240
Db	181	QTRCIRAPPNKGQVVLSSIEGRVAVELDPSPEVQKKYAFKCHRLKENNIEQIYPNVAI	240
Qy	241	SFNNIHTFATGSGDGFVNWDENKRLCQFHRYPTTSIASLAFSNDGTTLATASSYWE	300
Db	241	SFNNIHTFATGSGDGFVNWDENKRLCQFHRYPTTSIASLAFSNDGTTLATASSYWE	300
Qy	301	MDTDEHPEDGIFIRQVTDARTKPKSPCT 328	
Db	301	MDTDEHPEDGIFIRQVTDARTKPKSPCT 328	

RESULT 2
US-09-247-155-157

Sequence 157, Application US/09247155A
 Patent No. 6312922
 GENERAL INFORMATION:
 APPLICANT: Dumas Milne Edwards, Jean-Baptiste
 APPLICANT: Duclert, Aymeric
 APPLICANT: Bougueleret, Lydie
 TITLE OF INVENTION: Complementary DNAs
 FILE REFERENCE: GENSET.021A
 CURRENT FILING DATE: 1999-02-09
 EARLIER FILING DATE: 1998-02-09
 EARLIER FILING DATE: 1998-02-09
 EARLIER FILING DATE: 1998-04-13
 EARLIER FILING DATE: 1998-08-10
 EARLIER FILING DATE: 1998-08-10
 EARLIER FILING DATE: 1998-10-04
 NUMBER OF SEQ ID NOS: 182
 SOFTWARE: Patent.pm
 SEQ ID NO 157
 LENGTH: 59
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: SIGNAL
 LOCATION: -22...-1
 US-09-247-155-157

Query Match 14.6%; Score 258; DB 4; Length 59;
 Best Local Similarity 100.0%; Pred. No. 4.6e-21;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 MELKYQHTGAVIDCAFYDPTAWSGGLDHLQKXHDNLNTDQENLVGT 94
 DB 1 MELKYQHTGAVIDCAFYDPTAWSGGLDHLQKXHDNLNTDQENLVGT 46

RESULT 3

US-09-291-170A-13
 Sequence 13, Application US/09291170A
 Patent No. 6410687

GENERAL INFORMATION:
 APPLICANT: Vale, Ronald D.
 APPLICANT: Hartman, James J.
 APPLICANT: The Regents of the University of California
 TITLE OF INVENTION: Assays for the Detection of Microtubule
 TITLE OF INVENTION: Depolymerization Inhibitors
 FILE REFERENCE: 18557B-000510US
 CURRENT APPLICATION NUMBER: US/09/291,170A
 CURRENT FILING DATE: 1999-04-13
 PRIOR FILING DATE: 1998-04-14
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 13
 LENGTH: 251
 TYPE: PRT
 ORGANISM: Thermomonospora curvata
 FEATURE:
 OTHER INFORMATION: putative serine/threonine kinase Pkwa WD40 repeat
 OTHER INFORMATION: region

US-09-291-170A-13
 Query Match 11.1%; Score 197.5; DB 4; Length 251;
 Best Local Similarity 27.7%; Pred. No. 2.1e-13;
 Matches 70; Conservative 48; Mismatches 112; Indels 23; Gaps 10;

QY 15 DGISVVKFSPNTSOFLLVSSWDTSVRLYDVPANSMRLKYQ-HTGAVLDCAFY-DPTHAWS 72
 DB 14 DWVRAVAFSPD-GALLASGSDATVRLMDVAAEERAVFEGHTHYVLDIAFSPDGMVAS 72
 QY 73 GGLDHLQKXHDNLNTDQENLV-GTHDAPIRCVEYCPVNVVVTGSDQTVKLDWPTPCNA 131

Db 73 GSRDGTARLWNVATGTEHAHLKGHDTYVYAVAFSPDGMVASGSRDGTIRLMDVATGKER 132
 QY 132 GTFSQP-EKYVTLVSVDG-RLIVGTAGRRVLVMDLRNMGYVQORRESSLKYQTRCIRAPP 189
 Db 133 DVLQAPAEVNVSLAFSPDGMVHSGSDSTVHLMDVASGEALH-----TFEGHTDWRVA 187
 QY 190 -NKQGYVLSIEGRVAVEYLDPSPEVQKKYAFKCHRLKENNIEQIYVNAISFHNHT 248
 Db 188 FSPDGMVHSGSDSTVHLMDVAAQEE-----HTTLEGHTE---PVHSAFHPGEGT 236
 QY 249 FATGSDGDFVNIW 261
 Db 237 LASASEDGTIRIW 249

RESULT 4

US-09-724-884-13
 Sequence 13, Application US/09724884
 Patent No. 6429304

GENERAL INFORMATION:
 APPLICANT: Vale, Ronald D.
 APPLICANT: Hartman, James J.
 APPLICANT: The Regents of the University of California
 TITLE OF INVENTION: Assays for the Detection of Microtubule
 TITLE OF INVENTION: Depolymerization Inhibitors
 FILE REFERENCE: 18557B-000510US
 CURRENT APPLICATION NUMBER: US/09/724,884
 CURRENT FILING DATE: 2000-11-28
 PRIOR FILING DATE: 09/291,170
 PRIOR FILING DATE: 1999-04-13
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 13
 LENGTH: 251
 TYPE: PRT
 ORGANISM: Thermomonospora curvata
 FEATURE:
 OTHER INFORMATION: putative serine/threonine kinase Pkwa WD40 repeat
 OTHER INFORMATION: region

US-09-724-884-13
 Query Match 11.1%; Score 197.5; DB 4; Length 251;
 Best Local Similarity 27.7%; Pred. No. 2.1e-13;
 Matches 70; Conservative 48; Mismatches 112; Indels 23; Gaps 10;

QY 15 DGISVVKFSPNTSOFLLVSSWDTSVRLYDVPANSMRLKYQ-HTGAVLDCAFY-DPTHAWS 72
 Db 14 DWVRAVAFSPD-GALLASGSDATVRLMDVAAEERAVFEGHTHYVLDIAFSPDGMVAS 72
 QY 73 GGLDHLQKXHDNLNTDQENLV-GTHDAPIRCVEYCPVNVVVTGSDQTVKLDWPTPCNA 131
 Db 73 GSRDGTARLWNVATGTEHAHLKGHDTYVYAVAFSPDGMVASGSRDGTIRLMDVATGKER 132
 QY 132 GTFSQP-EKYVTLVSVDG-RLIVGTAGRRVLVMDLRNMGYVQORRESSLKYQTRCIRAPP 189
 Db 133 DVLQAPAEVNVSLAFSPDGMVHSGSDSTVHLMDVASGEALH-----TFEGHTDWRVA 187
 QY 190 -NKQGYVLSIEGRVAVEYLDPSPEVQKKYAFKCHRLKENNIEQIYVNAISFHNHT 248
 Db 188 FSPDGMVHSGSDSTVHLMDVAAQEE-----HTTLEGHTE---PVHSAFHPGEGT 236
 QY 249 FATGSDGDFVNIW 261
 Db 237 LASASEDGTIRIW 249

RESULT 5

US-09-063-743-1
 Sequence 1, Application US/09063743
 Patent No. 6242214

GENERAL INFORMATION:
 APPLICANT: Bandman, Olga

APPLICANT: Lal, Preeti
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Patterson, Chandra
TITLE OF INVENTION: HUMAN GTPASE-ASSOCIATED PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA: US/09/063,743
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0508 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNGN0701
CLONE: 059953
US-09-063-743-1

Query Match 10.0%; Score 176.5; DB 3; Length 375;
Best Local Similarity 25.8%; Pred. No. 8.2e-11;
Matches 80; Conservative 51; Mismatches 142; Indels 37; Gaps 13;
QY 14 EDGISSVKFSPNTSQELLVSSWDTSVRLYDVPA---NSMRLKYQHTGAVLDCAF-YDPHT 69
Db 66 EGEVYCKFHNGST-LASAGFDRLLMNVYDCDNYATLK-GHSGAVMELHNTDGS 123
QY 70 AWSGGLDHLQKMHDLNTDQ--ENLVGTHDAPIRCVEYCEPVNMVMTGSDQTVKLDPR 127
Db 124 LFSASTDKTVAVMDSETGERVKRLKHTSFVNSCYPARRGPQLVCTGSDGDTGLWDIRK 183
QY 128 PCNAGTFSQPEKV--YTLVSGRLVGTAGRRVLVMDLRNMGYVQQRRESSLKYQTR-- 183
Db 184 KAAIQFTQNTYQVLAVTFNDTSQIIISGGINDIKVMDL-----RQNKLYTMRGH 234
QY 184 ----CIRAFPNKQGVLSIEGRVAVVEYLDPSPEVQKKYAFKCHRLKENNIEQIYPVA 239
Db 289 LRCSWSPDGSKIAAGSADREVYVMDTTSRRILYKLPCHAGSINEVAFHDEPTIISASD 348
QY 297 ---YVMEDD 303
Db 349 KRLYMGEFSE 358

RESULT 6
US-09-590-540-1
Sequence 1, Application US/09590540
Patent No. 6410267
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Lal, Preeti
Guegler, Karl J.
Corley, Neil C.
Patterson, Chandra
TITLE OF INVENTION: HUMAN GTPASE-ASSOCIATED PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Genomics, Inc.
STREET: 3160 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 3.0
CURRENT APPLICATION DATA: US/09/590,540
FILING DATE: 08-Jun-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/063,743
FILING DATE: April 21, 1998
ATTORNEY/AGENT INFORMATION:
NAME: David G. Streeter
REGISTRATION NUMBER: 43,168
REFERENCE/DOCKET NUMBER: PF-0508-1 DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNGN0701
CLONE: 059953
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-590-540-1

Query Match 10.0%; Score 176.5; DB 4; Length 375;
Best Local Similarity 25.8%; Pred. No. 8.2e-11;
Matches 80; Conservative 51; Mismatches 142; Indels 37; Gaps 13;
QY 14 EDGISSVKFSPNTSQELLVSSWDTSVRLYDVPA---NSMRLKYQHTGAVLDCAF-YDPHT 69
Db 66 EGEVYCKFHNGST-LASAGFDRLLMNVYDCDNYATLK-GHSGAVMELHNTDGS 123
QY 70 AWSGGLDHLQKMHDLNTDQ--ENLVGTHDAPIRCVEYCEPVNMVMTGSDQTVKLDPR 127
Db 124 LFSASTDKTVAVMDSETGERVKRLKHTSFVNSCYPARRGPQLVCTGSDGDTGLWDIRK 183
QY 128 PCNAGTFSQPEKV--YTLVSGRLVGTAGRRVLVMDLRNMGYVQQRRESSLKYQTR-- 183
Db 184 KAAIQFTQNTYQVLAVTFNDTSQIIISGGINDIKVMDL-----RQNKLYTMRGH 234
QY 184 ----CIRAFPNKQGVLSIEGRVAVVEYLDPSPEVQKKYAFKCHRLKENNIEQIYPVA 239
Db 235 ADSVTGLSLSEGSYLLSNAMDN-TVRVMDVRFPAPKE---RCVKIFQGNVHN-FEKNL 288
QY 240 I--SFHNIHTFATGSDGPFVNIWDPFNKKELCQPHRYPTSIASLAFSNDGTTLAISS- 296

```
Db 289 LRCSMSPGSKIAAGSADRFFVVMDDTSRRILYKLPCHAGSINEVAHPHPDEPIIASSD 348
Qy 297 ---YVEMDD 303
Db 349 KRLYMGEPSE 358

RESULT 7
US-09-291-170A-12
; Sequence 12, Application US/09291170A
; Patent No. 6410687
; GENERAL INFORMATION:
; APPLICANT: Vale, Ronald D.
; APPLICANT: Hartman, James J.
; TITLE OF INVENTION: Assays for the Detection of Microtubule
; TITLE OF INVENTION: Depolymerization Inhibitors
; FILE REFERENCE: 18557B-000510US
; CURRENT FILING DATE: 1999-04-13
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: TFIID WD40 repeat region
US-09-291-170A-12

Query Match          9.6%; Score 169.5; DB 4; Length 250;
Best Local Similarity 21.5%; Pred. No. 2.5e-10;
Matches 69; Conservative 44; Mismatches 101; Indels 107; Gaps 9;

Qy 4 SNEFKLNQPPEDGIGSSVKFSPNTSQFLVSSWMTSVRLYDVDPANSMRLKYQ-HTGAVLDC 62
Db 3 ASELKILYGHSGPYVGASFSFDRN-YLLSSSEDGTVRLMSLQTFCLVGYKGHNYPVMDT 61
Qy 63 AFYDPTHAM-SGGLDHQKMDLNTDQE-NLVGTHDAPIRCEYCEPVNMVVTGSDQTV 120
Db 62 QFSPGYGYFVSGGHRVARLWATDHYQPLRIFAGHLADVNCITRPHNSNYVATGSADRTV 121
Qy 121 KLMDPRTPCNAGTFSQPEKVTLSVSGDRLVGTAGRRVLWDLRNMGYVQORRESSLKY 180
Db 122 RLMD-----VLNGN----- 130
Qy 181 QTRCIRAFPNKQGYVLLSIEGRVAVEYLDPSPEVQKKYAFKCHRLKENNIEQIYPVNAI 240
Db 131 ---CVRIFTGHKG-----PIHSL 145
Qy 241 SFNHNFTATGSGDGFVNIWDPFNKRLCQFHYPTSIASLAFSNDGTTLAIASSYME 300
Db 146 TFSNGRFLATGATDGRVLLWDIGHGLMVGLKGTDTVCSLFRSDRDEILASGS-----200
Qy 301 MDDTEHPEDGIFIRQVTDDET 321
Db 201 MDMTVRLMDA--IKAFEDLET 219

RESULT 9
US-08-190-802A-38
; Sequence 38, Application US/08190802A
; Patent No. 5519003
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,802A
; FILING DATE: 01-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 8600-0139
; TELECOMMUNICATION INFORMATION:
```

TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 38:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 340 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: G-Beta 1 bovine, Fig. 21
 US-08-190-802A-38

Query Match 9.6%; Score 169.5; DB 1; Length 340;
 Best Local Similarity 24.8%; Pred. No. 4.1e-10;
 Matches 73; Conservative 43; Mismatches 127; Indels 51; Gaps 14;

QY 26 TSQFLVSS-----WD--TSVRLYDVPANS---MRLKYQHTGAVLDCAFYDPHTAWS 72
 DB 65 TDSRLVSAQDGKLIWDSYTTNKVHAIFLRSSWMTCAYPGNYVAC-----114
 QY 73 GGLDHQLKMDLNTDQNL-----VGTDAPIRCVCEPVNVMVTGSDQTVKLDPRPT 127
 DB 115 GGLDNICSYNLTREGNVRVRELAGHTGYLSCCFRLDD-NQIVTSSGDTTCALWDIET 173
 QY 128 PCNAGTFS-OPEKVVYTLVSVDG-RLIV-GTAGRRVLVMDLRNMGYVQOR---RESSLKQY 181
 DB 174 GQOTTFTTGTGDVMSLSLAPDTRLFVSGACDASAKLMDVRE-GMCRQFTTGHESDI---229
 QY 182 TRCIRAFPNKQGVLSIEGRVAVELDSPVEQKKYAFKCHRLKENNIEQIYPVNAIS 241
 DB 230 -NAICFPNGNAFATGSDDATCRLFDLRADQELMTYSH-----DNI--ICGITSVS 277
 QY 242 FHNHNTATGSGDFNINWDPFNKRLCOFHRYPTASIASLAPNSDGTTLAIAS 295
 DB 278 FSKSGRLLAGYDDFNCVWDALKADRAVLGHNDNRVSLGVTDDGMAVATCS 331

RESULT 10

US-08-477-346-38
 Sequence 38, Application US/08477346
 Patent No. 6262023
 GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 TITLE OF INVENTION: Thereof
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477,346
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/487,072
 FILING DATE: 07-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.20
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 INFORMATION FOR SEQ ID NO: 38:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 340 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: G-Beta 1 bovine, Fig. 21
 US-08-477-346-38

Query Match 9.6%; Score 169.5; DB 3; Length 340;
 Best Local Similarity 24.8%; Pred. No. 4.1e-10;
 Matches 73; Conservative 43; Mismatches 127; Indels 51; Gaps 14;

QY 26 TSQFLVSS-----WD--TSVRLYDVPANS---MRLKYQHTGAVLDCAFYDPHTAWS 72
 DB 65 TDSRLVSAQDGKLIWDSYTTNKVHAIFLRSSWMTCAYPGNYVAC-----114
 QY 73 GGLDHQLKMDLNTDQNL-----VGTDAPIRCVCEPVNVMVTGSDQTVKLDPRPT 127
 DB 115 GGLDNICSYNLTREGNVRVRELAGHTGYLSCCFRLDD-NQIVTSSGDTTCALWDIET 173
 QY 128 PCNAGTFS-OPEKVVYTLVSVDG-RLIV-GTAGRRVLVMDLRNMGYVQOR---RESSLKQY 181
 DB 174 GQOTTFTTGTGDVMSLSLAPDTRLFVSGACDASAKLMDVRE-GMCRQFTTGHESDI---229
 QY 182 TRCIRAFPNKQGVLSIEGRVAVELDSPVEQKKYAFKCHRLKENNIEQIYPVNAIS 241
 DB 230 -NAICFPNGNAFATGSDDATCRLFDLRADQELMTYSH-----DNI--ICGITSVS 277
 QY 242 FHNHNTATGSGDFNINWDPFNKRLCOFHRYPTASIASLAPNSDGTTLAIAS 295
 DB 278 FSKSGRLLAGYDDFNCVWDALKADRAVLGHNDNRVSLGVTDDGMAVATCS 331

RESULT 11

US-08-473-089-38
 Sequence 38, Application US/08473089
 Patent No. 6342368
 GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 TITLE OF INVENTION: Thereof
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/473,089
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.22
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: G-Beta 1 bovine, Fig. 21

US-08-473-089-38

Query Match 9.6%; Score 169.5; DB 4; Length 340;
Best Local Similarity 24.8%; Pred. No. 4.1e-10;
Matches 73; Conservative 43; Mismatches 127; Indels 51; Gaps 14;
QY 26 TSQFLVSS-----WD--TSVRLYDVPPANS---MRLKYQHTGAVLDCAFYDPTTHAWS 72
DB 65 TDSRLVSSASQDGKLIWDSYTTNKVHAIPLRSSWMTCAVAPSGNVAC-----114
QY 73 GGLDHQLKXHLNLTDOENL-----VGTHDAPIRCEVCEPVNVMVTGSDQTVKLDPRPT 127
DB 115 GGLDNICSIYNLKTREGNVRVSRRELAGHTGYLSCCRFLDD-NQIVTSSGDTTCALWDIET 173
QY 128 PCNAGTFS-QPEKVYTLVSQD-RLIV-GTAGRRVLVWDLRNMGYVQOR---RESSLKYQ 181
DB 174 GQQTTFGTGHTGDVMSLSLAPDTRLFVSGACDASAKLMDVRE-GMCRQTFTHESDI---229
QY 182 TRCIRAFPNKQGYVLSISIEGRVAVEYLDPSPEVQKKYAFKCHRLKNNIEQIYPVNAIS 241
DB 230 -NAICFFPNAGFATGSDATCRFLDLRADELMYSH-----DNI--ICGITSVS 277
QY 242 FHNHNTFATGSGDFVNIWDPFNKKRLCOQHRYPTSIASLAFSNDGTTLAIAS 295
DB 278 FSKSGRLLLAGYDDFNCNVMDALKADRAVLGHNDNRVSCGLVTDGMAVATGS 331

RESULT 12

US-08-487-072A-38
Sequence 38, Application US/08487072A
Patent No. 6423684
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,072A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid

TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: G-Beta 1 bovine, Fig. 21

US-08-487-072A-38

Query Match 9.6%; Score 169.5; DB 4; Length 340;
Best Local Similarity 24.8%; Pred. No. 4.1e-10;
Matches 73; Conservative 43; Mismatches 127; Indels 51; Gaps 14;
QY 26 TSQFLVSS-----WD--TSVRLYDVPPANS---MRLKYQHTGAVLDCAFYDPTTHAWS 72
DB 65 TDSRLVSSASQDGKLIWDSYTTNKVHAIPLRSSWMTCAVAPSGNVAC-----114
QY 73 GGLDHQLKXHLNLTDOENL-----VGTHDAPIRCEVCEPVNVMVTGSDQTVKLDPRPT 127
DB 115 GGLDNICSIYNLKTREGNVRVSRRELAGHTGYLSCCRFLDD-NQIVTSSGDTTCALWDIET 173
QY 128 PCNAGTFS-QPEKVYTLVSQD-RLIV-GTAGRRVLVWDLRNMGYVQOR---RESSLKYQ 181
DB 174 GQQTTFGTGHTGDVMSLSLAPDTRLFVSGACDASAKLMDVRE-GMCRQTFTHESDI---229
QY 182 TRCIRAFPNKQGYVLSISIEGRVAVEYLDPSPEVQKKYAFKCHRLKNNIEQIYPVNAIS 241
DB 230 -NAICFFPNAGFATGSDATCRFLDLRADELMYSH-----DNI--ICGITSVS 277
QY 242 FHNHNTFATGSGDFVNIWDPFNKKRLCOQHRYPTSIASLAFSNDGTTLAIAS 295
DB 278 FSKSGRLLLAGYDDFNCNVMDALKADRAVLGHNDNRVSCGLVTDGMAVATGS 331

RESULT 13

US-09-245-039-1
Sequence 1, Application US/09245039
Patent No. 6555522
GENERAL INFORMATION:
APPLICANT: Iyengar, Srinivas R.V.
APPLICANT: Weng, Gezhi
APPLICANT: Chen, Yibang
APPLICANT: Weinstein, Harel
APPLICANT: Buck, Elizabeth
TITLE OF INVENTION: PEPTIDES AND OTHER SMALL MOLECULES DERIVED FROM REGIONS
TITLE OF INVENTION: OF INTERACTING PROTEINS AND USES THEREOF
FILE REFERENCE: 6923-074
CURRENT APPLICATION NUMBER: US/09/245,039
CURRENT FILING DATE: 1999-02-05
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 340
TYPE: PRT
ORGANISM: Homo sapiens
US-09-245-039-1
Query Match 9.6%; Score 169.5; DB 4; Length 340;
Best Local Similarity 24.8%; Pred. No. 4.1e-10;
Matches 73; Conservative 43; Mismatches 127; Indels 51; Gaps 14;
QY 26 TSQFLVSS-----WD--TSVRLYDVPPANS---MRLKYQHTGAVLDCAFYDPTTHAWS 72
DB 65 TDSRLVSSASQDGKLIWDSYTTNKVHAIPLRSSWMTCAVAPSGNVAC-----114
QY 73 GGLDHQLKXHLNLTDOENL-----VGTHDAPIRCEVCEPVNVMVTGSDQTVKLDPRPT 127
DB 115 GGLDNICSIYNLKTREGNVRVSRRELAGHTGYLSCCRFLDD-NQIVTSSGDTTCALWDIET 173
QY 128 PCNAGTFS-QPEKVYTLVSQD-RLIV-GTAGRRVLVWDLRNMGYVQOR---RESSLKYQ 181
DB 174 GQQTTFGTGHTGDVMSLSLAPDTRLFVSGACDASAKLMDVRE-GMCRQTFTHESDI---229
QY 182 TRCIRAFPNKQGYVLSISIEGRVAVEYLDPSPEVQKKYAFKCHRLKNNIEQIYPVNAIS 241

[illegible]

```

QY 301 MDDTEHPDGIFIRQVTDAT 321
DB 635 MDMTVRIWDA--IKAFEDLET 653

RESULT 15
US-08-646-715-18
; Sequence 18, Application US/08646715
; Patent No. 5637886
; GENERAL INFORMATION:
; APPLICANT: Tjian, Robert
; APPLICANT: Comal, Lucio
; APPLICANT: Dynlacht, Brian D.
; APPLICANT: Hoey, Timothy
; APPLICANT: Ruppert, Siegfried
; APPLICANT: Tanese, Naoko
; APPLICANT: Wang, Edith
; APPLICANT: Weinzierl, Robert O.J.
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,715
; FILING DATE: 09-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/188,582
; FILING DATE: 28-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 704 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-646-715-18

Query Match 9.6%; Score 169.5; DB 1; Length 704;
Best Local Similarity 21.5%; Pred. No. 1.3e-09;
Matches 69; Conservative 44; Mismatches 101; Indels 107; Gaps 9;

QY 4 SNEFKLNQPPDGISGVKFSPTNSOFLVLVSWSDTSVRLYDVPANSMRLKYQ-HTGAVLDC 62
DB 437 ASEKLTYLGHSGPVYGASFSPORN-YLLSSSDGTGTVRLWSLQTTCTCLGVYKGHNYPVWDT 495
QY 63 AFYDPTHAM-SGGLDHLQKHLNDTDOE-NLVGTHDAFIRCEVCPEVNNWVTGSDQTV 120

```



```

db 496 QFSPXYFVSGHQRVARLWATHYQPLRIFAGHLADVNCTRFHPNSNYVATGSADRTV 555
QY 121 KLMDPRTPCNAGTFSQPEKVYTLVSXGDLIVGTAGRRVLWDLRNMGVVQORRESSLKY 180
Db 556 RLD-----VLGN----- 564
QY 181 QTRCIRAFPNKQGYVLSIEGRVAVEYLDPSPEVQKKYAFKCHRLKENNIEQIYVNAI 240
Db 565 ---CVRIFTGHKG-----PIHSL 579
QY 241 SFHNHNTFATGSDGFVNINDPFPNKKRLCQPHRYPTSIASLAPNDGTTLAIASSYMYE 300
Db 580 TFSNGRFATGATDGRVLLNDIGHGLMVGELKHTDVCSLRFSRDEILASGS----- 634
QY 301 MDDTEHPEDGIFIRQVDAET 321
Db 635 MONTVRLWDA--IKAFEDLET 653

```

Search completed: November 12, 2003, 20:11:20
 Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 20:08:53 ; Search time 30 Seconds
(without alignments)
1877.792 Million cell updates/sec

Title: US-10-084-700-2

Perfect score: 1772
Sequence: 1 MTGSNEFKLNQPPEDGISSV.....DGIFIRQVDAETKPKSPCT 328

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA.*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1772	100.0	328	10	US-09-095-881-4
2	1772	100.0	328	14	US-10-084-700-2
3	1772	100.0	332	9	US-09-925-301-1301
4	1743	98.6	330	14	US-10-084-700-27
5	1743	98.4	326	14	US-10-084-700-29
6	538	30.4	365	14	US-10-084-700-32
7	511.5	28.9	352	14	US-10-084-700-31
8	509	28.7	368	14	US-10-084-700-30
9	509	28.7	413	9	US-09-925-302-712
10	356	20.1	341	14	US-10-084-700-28
11	258	14.6	59	12	US-09-903-130-157
12	197.5	11.1	100	9	US-09-925-299-1045
13	197.5	11.1	100	11	US-09-925-299-1045
14	197.5	11.1	742	14	US-10-077-111-11
15	176.5	10.0	375	14	US-10-119-932-1

16	176	9.9	1356	14	US-10-077-111-10	Sequence 10, Appli
17	175.5	9.9	486	15	US-10-274-525-3	Sequence 3, Appli
18	169.5	9.6	340	10	US-09-226-248B-2	Sequence 2, Appli
19	169.5	9.6	340	12	US-10-320-778-1	Sequence 1, Appli
20	166.5	9.4	281	12	US-10-245-618-20	Sequence 20, Appli
21	166.5	9.4	540	10	US-09-213-888-7	Sequence 7, Appli
22	166.5	9.4	540	10	US-09-213-888-10	Sequence 10, Appli
23	166.5	9.4	540	10	US-09-328-877A-7	Sequence 7, Appli
24	166.5	9.4	540	10	US-09-328-877A-10	Sequence 10, Appli
25	166.5	9.4	540	12	US-10-245-618-14	Sequence 14, Appli
26	166.5	9.4	545	10	US-09-213-888-6	Sequence 6, Appli
27	166.5	9.4	545	10	US-09-328-877A-6	Sequence 6, Appli
28	166.5	9.4	553	10	US-09-213-888-5	Sequence 5, Appli
29	166.5	9.4	553	10	US-09-328-877A-5	Sequence 5, Appli
30	166.5	9.4	553	12	US-10-245-618-8	Sequence 8, Appli
31	166.5	9.4	559	10	US-09-213-888-9	Sequence 9, Appli
32	166.5	9.4	559	10	US-09-328-877A-9	Sequence 9, Appli
33	166.5	9.4	589	10	US-09-213-888-8	Sequence 8, Appli
34	166.5	9.4	589	10	US-09-328-877A-8	Sequence 8, Appli
35	166.5	9.4	592	10	US-10-245-618-18	Sequence 18, Appli
36	166.5	9.4	592	10	US-09-213-888-4	Sequence 4, Appli
37	166.5	9.4	626	10	US-09-328-877A-4	Sequence 4, Appli
38	166.5	9.4	626	10	US-09-213-888-21	Sequence 21, Appli
39	166.5	9.4	627	10	US-09-328-877A-21	Sequence 21, Appli
40	166.5	9.4	627	10	US-09-213-888-3	Sequence 3, Appli
41	166.5	9.4	627	10	US-09-328-877A-3	Sequence 3, Appli
42	166.5	9.4	627	12	US-10-245-618-6	Sequence 6, Appli
43	166.5	9.4	666	10	US-09-213-888-27	Sequence 27, Appli
44	166.5	9.4	666	10	US-09-328-877A-27	Sequence 27, Appli
45	166.5	9.4	669	10	US-09-213-888-25	Sequence 25, Appli

ALIGNMENTS

RESULT 1

US-09-095-881-4
; Sequence 4, Application US/09095881
; Patent No. US20020123042A1
; GENERAL INFORMATION:
; APPLICANT: Seeley, Todd
; TITLE OF INVENTION: DETECTION OF LOSS OF THE WILD-TYPE
; FILE REFERENCE: 1405.003 / 200130.437
; CURRENT APPLICATION NUMBER: US/09/095,881
; CURRENT FILING DATE: 1998-06-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-095-881-4

Query Match 100.0%; Score 1772; DB 10; Length 328;
Best Local Similarity 100.0%; Pred. No. 6e-180;
Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MTGSNEFKLNQPPEDGISSVKFSPNTSQFLVSSWDTSVRLVDYVPANSMRLKYQHTGAVL	60
DB	1	MTGSNEFKLNQPPEDGISSVKFSPNTSQFLVSSWDTSVRLVDYVPANSMRLKYQHTGAVL	60
QY	61	DCAFYDPTAWSGGLDHLQKXHDLTQENLVGTHDAPIRCVCEYCPENVVMTGSDQTV	120
DB	61	DCAFYDPTAWSGGLDHLQKXHDLTQENLVGTHDAPIRCVCEYCPENVVMTGSDQTV	120
QY	121	KLWDPRTFCNAGTSPQEKVYTLVSQDRLIVGTAGRRVLVWDLRNNGYVQQRRESLKY	180
DB	121	KLWDPRTFCNAGTSPQEKVYTLVSQDRLIVGTAGRRVLVWDLRNNGYVQQRRESLKY	180
QY	181	QTRCIRAFPNKQGVLLSSIEGRVAVELDPSPVQKKYAFKCHRLKNNIEQIYPYNAI	240
DB	181	QTRCIRAFPNKQGVLLSSIEGRVAVELDPSPVQKKYAFKCHRLKNNIEQIYPYNAI	240

QY 241 SFNHNHTFATGSDGDFVNIWDPFNKRLCOFHRYPTSIASLAFSNDGTTTIAIASSYMYE 300
DB 241 SFNHNHTFATGSDGDFVNIWDPFNKRLCOFHRYPTSIASLAFSNDGTTTIAIASSYMYE 300
QY 301 MDDTEHPEDGIFIRQVTDATKPKSPCT 328
DB 301 MDDTEHPEDGIFIRQVTDATKPKSPCT 328

RESULT 2

US-10-084-700-2
; Sequence 2, Application US/10084700
; Publication No. US20020160403A1
; GENERAL INFORMATION:
; APPLICANT: Seeley, Todd
; TITLE OF INVENTION: hUBB3 GENE INVOLVED IN HUMAN CANCERS
; FILE REFERENCE: PP-01406.004/200130.43801
; CURRENT APPLICATION NUMBER: US/10/084,700
; CURRENT FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-084-700-2

Query Match 100.0%; Score 1772; DB 14; Length 328;
Best Local Similarity 100.0%; Pred. No. 6e-180;
Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGSNEFKLNQPPEDGISSVKFSPNTSQFLVSSWSDTSVRLYDVPANSMRLKYQHTGAVL 60
DB 1 MTGSNEFKLNQPPEDGISSVKFSPNTSQFLVSSWSDTSVRLYDVPANSMRLKYQHTGAVL 60
QY 61 DCAFYDPTHAMSGGLDHLQKMDLNTQENLVGTHDAPIRCVCEYCPVNVMTGSDQTV 120
DB 61 DCAFYDPTHAMSGGLDHLQKMDLNTQENLVGTHDAPIRCVCEYCPVNVMTGSDQTV 120
QY 121 KLMDPRTPCNAGTFSQPEKVVYTLVSGDRLIVGTAGRRVLVMDLRNMGYVQQRRESSLKY 180
DB 121 KLMDPRTPCNAGTFSQPEKVVYTLVSGDRLIVGTAGRRVLVMDLRNMGYVQQRRESSLKY 180
QY 181 QTRCIRAFPNKQGYVLSSIEGRVAVVEYLDPSPEVQKKYAFKCHRLKENNIEQIYPVNAI 240
DB 181 QTRCIRAFPNKQGYVLSSIEGRVAVVEYLDPSPEVQKKYAFKCHRLKENNIEQIYPVNAI 240
QY 241 SFNHNHTFATGSDGDFVNIWDPFNKRLCOFHRYPTSIASLAFSNDGTTTIAIASSYMYE 300
DB 241 SFNHNHTFATGSDGDFVNIWDPFNKRLCOFHRYPTSIASLAFSNDGTTTIAIASSYMYE 300
QY 301 MDDTEHPEDGIFIRQVTDATKPKSPCT 328
DB 301 MDDTEHPEDGIFIRQVTDATKPKSPCT 328

RESULT 3

US-09-925-301-1301
; Sequence 1301, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1301
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1301

Query Match 100.0%; Score 1772; DB 9; Length 332;
Best Local Similarity 100.0%; Pred. No. 6.1e-180;
Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGSNEFKLNQPPEDGISSVKFSPNTSQFLVSSWSDTSVRLYDVPANSMRLKYQHTGAVL 60
DB 5 MTGSNEFKLNQPPEDGISSVKFSPNTSQFLVSSWSDTSVRLYDVPANSMRLKYQHTGAVL 64
QY 61 DCAFYDPTHAMSGGLDHLQKMDLNTQENLVGTHDAPIRCVCEYCPVNVMTGSDQTV 120
DB 65 DCAFYDPTHAMSGGLDHLQKMDLNTQENLVGTHDAPIRCVCEYCPVNVMTGSDQTV 124
QY 121 KLMDPRTPCNAGTFSQPEKVVYTLVSGDRLIVGTAGRRVLVMDLRNMGYVQQRRESSLKY 180
DB 125 KLMDPRTPCNAGTFSQPEKVVYTLVSGDRLIVGTAGRRVLVMDLRNMGYVQQRRESSLKY 184
QY 181 QTRCIRAFPNKQGYVLSSIEGRVAVVEYLDPSPEVQKKYAFKCHRLKENNIEQIYPVNAI 240
DB 185 QTRCIRAFPNKQGYVLSSIEGRVAVVEYLDPSPEVQKKYAFKCHRLKENNIEQIYPVNAI 244
QY 241 SFNHNHTFATGSDGDFVNIWDPFNKRLCOFHRYPTSIASLAFSNDGTTTIAIASSYMYE 300
DB 245 SFNHNHTFATGSDGDFVNIWDPFNKRLCOFHRYPTSIASLAFSNDGTTTIAIASSYMYE 304
QY 301 MDDTEHPEDGIFIRQVTDATKPKSPCT 328
DB 305 MDDTEHPEDGIFIRQVTDATKPKSPCT 332

RESULT 4

US-10-084-700-27
; Sequence 27, Application US/10084700
; Publication No. US20020160403A1
; GENERAL INFORMATION:
; APPLICANT: Seeley, Todd
; TITLE OF INVENTION: hUBB3 GENE INVOLVED IN HUMAN CANCERS
; FILE REFERENCE: PP-01406.004/200130.43801
; CURRENT APPLICATION NUMBER: US/10/084,700
; CURRENT FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-084-700-27

Query Match 98.6%; Score 1747; DB 14; Length 330;
Best Local Similarity 100.0%; Pred. No. 2.8e-177;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGSNEFKLNQPPEDGISSVKFSPNTSQFLVSSWSDTSVRLYDVPANSMRLKYQHTGAVL 60
DB 1 MTGSNEFKLNQPPEDGISSVKFSPNTSQFLVSSWSDTSVRLYDVPANSMRLKYQHTGAVL 60
QY 61 DCAFYDPTHAMSGGLDHLQKMDLNTQENLVGTHDAPIRCVCEYCPVNVMTGSDQTV 120
DB 61 DCAFYDPTHAMSGGLDHLQKMDLNTQENLVGTHDAPIRCVCEYCPVNVMTGSDQTV 120
QY 121 KLMDPRTPCNAGTFSQPEKVVYTLVSGDRLIVGTAGRRVLVMDLRNMGYVQQRRESSLKY 180
DB 121 KLMDPRTPCNAGTFSQPEKVVYTLVSGDRLIVGTAGRRVLVMDLRNMGYVQQRRESSLKY 180
QY 181 QTRCIRAFPNKQGYVLSSIEGRVAVVEYLDPSPEVQKKYAFKCHRLKENNIEQIYPVNAI 240
DB 181 QTRCIRAFPNKQGYVLSSIEGRVAVVEYLDPSPEVQKKYAFKCHRLKENNIEQIYPVNAI 240

QY 241 SFNHNHTFATGSGDGFVNIWDPFNKRLCQFHRYPTSIASLAFSNDGTTLAIASSYWE 300
Db 241 SFNHNHTFATGSGDGFVNIWDPFNKRLCQFHRYPTSIASLAFSNDGTTLAIASSYWE 300
QY 301 MDDTEHPEDGIFIRQVTDATKPK 324
Db 301 MDDTEHPEDGIFIRQVTDATKPK 324
RESULT 5
US-10-084-700-29
; Sequence 29, Application US/10084700
; Publication No. US20020160403A1
; GENERAL INFORMATION:
; APPLICANT: Seeley, Todd
; TITLE OF INVENTION: hUBU3 GENE INVOLVED IN HUMAN CANCERS
; FILE REFERENCE: PP-01406.004/200130.438D1
; CURRENT APPLICATION NUMBER: US/10/084,700
; CURRENT FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-084-700-29

Query Match 98.4%; Score 1743; DB 14; Length 326;
Best Local Similarity 99.7%; Pred. No. 7.2e-177;
Matches 324; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MTGSNEFKLQPPEDGIGSSVKFSNTSOFLLVSSWSDTSVRLYDVPANSMELKYQHTGAVL 60
Db 1 MTGSNEFKLQPPEDGIGSSVKFSNTSOFLLVSSWSDTSVRLYDVPANSMELKYQHTGAVL 60
QY 61 DCAFYDPTHAWSGGLDHLQKMHDLNTDQENLVGTHDAPIRCVCEYCPVNVMTGSDQTV 120
Db 61 DCAFYDPTHAWSGGLDHLQKMHDLNTDQENLVGTHDAPIRCVCEYCPVNVMTGSDQTV 120
QY 121 KLMDPRTPCNAGTFSQBEKVYTLVSQDRLIVGTAGRRVLWDLRNNGYVQQRRESLKY 180
Db 121 KLMDPRTPCNAGTFSQBEKVYTLVSQDRLIVGTAGRRVLWDLRNNGYVQQRRESLKY 180
QY 181 QTCIRAFPNKQGVLSISIEGRVAVELDPSPEVOKKYAFKCHRLKENNIEIYVNAI 240
Db 181 QTCIRAFPNKQGVLSISIEGRVAVELDPSPEVOKKYAFKCHRLKENNIEIYVNAI 240
QY 241 SFNHNHTFATGSGDGFVNIWDPFNKRLCQFHRYPTSIASLAFSNDGTTLAIASSYWE 300
Db 241 SFNHNHTFATGSGDGFVNIWDPFNKRLCQFHRYPTSIASLAFSNDGTTLAIASSYWE 300
QY 301 MDDTEHPEDGIFIRQVTDATKPKS 325
Db 301 MDDTEHPEDGIFIRQVTDATKPKS 325

RESULT 6
US-10-084-700-32
; Sequence 32, Application US/10084700
; Publication No. US20020160403A1
; GENERAL INFORMATION:
; APPLICANT: Seeley, Todd
; TITLE OF INVENTION: hUBU3 GENE INVOLVED IN HUMAN CANCERS
; FILE REFERENCE: PP-01406.004/200130.438D1
; CURRENT APPLICATION NUMBER: US/10/084,700
; CURRENT FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-084-700-32

Query Match 30.4%; Score 538; DB 14; Length 365;
Best Local Similarity 35.7%; Pred. No. 1.7e-48;
Matches 122; Conservative 59; Mismatches 135; Indels 26; Gaps 10;
QY 4 SNEFKLNQPPEDGIGSSVKFSNTSOFLLVSSWSDTSVRLYDVPANSMELKYQH-TGAVLDC 62
Db 26 ANDIVINSFAEDSISDIAFSPQDDPMFSASSMDKVRIMDVQGVQGRAQHESSSPVLC 85
QY 63 AFY--DPTHAWSGGLDHLQKMHDLNTDQENLVGTHDAPIRCVCEY--CPEVNV--MVTGSM 116
Db 86 TRWSNDGTKVASGGCDNALKLYDIASGQTQOIGHSAPIKVLRFVQCGSPNTCIVTGSW 145
QY 117 DOTVKLMDPRTPCNAGTFSQBEKVYTLVSQDRLIVGTAGRRVLWDLRNNGYVQQRRES 176
Db 146 DKTIKYMDRQPPQPVSTVMPERVSMQNKQSLVAVATAERHTAIINLANPTTIFKATIS 205
QY 177 SLKYQTRCTRAPFNKQGVLSISIEGRVAVELDPSPEVOKKK-YAFKCHRLK-----E 228
Db 206 PLKWQTRCVACNEADGAYGIVSGVGRCSIRYIDDG--MOKKSGFSFKCHRTNPAPGS 263
QY 229 NNIEQIYVNAISFHNHNTFATGSGDGFVNIWDPFNKRLCQFHRYPTSIASL---AFS 285
Db 264 NQSLVYVPSNIAFHPLYGTFTVGTAGDGTGFNFDKKNQHRHL---KGYPTLQASIPVCSFN 320
QY 286 NGDTTLAIASSYMY---EMDDTEHPEDGIFIRQVTDATKPK 324
Db 321 RNSGVFAVYALSDWHQGHMGNRPDPVNPVIRLHATTDEEVKEK 362

RESULT 7
US-10-084-700-31
; Sequence 31, Application US/10084700
; Publication No. US20020160403A1
; GENERAL INFORMATION:
; APPLICANT: Seeley, Todd
; TITLE OF INVENTION: hUBU3 GENE INVOLVED IN HUMAN CANCERS
; FILE REFERENCE: PP-01406.004/200130.438D1
; CURRENT APPLICATION NUMBER: US/10/084,700
; CURRENT FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-084-700-31

Query Match 28.9%; Score 511.5; DB 14; Length 352;
Best Local Similarity 34.8%; Pred. No. 1.1e-45;
Matches 115; Conservative 58; Mismatches 142; Indels 15; Gaps 7;
QY 6 EFKLNQPPEDGIGSSVKFSNTSOFLLVSSWSDTSVRLYDVPANSM---RLKYQHTGAVLDC 62
Db 22 DVEVAQPPEDSISDLAFSPQ-AEYLAASWSDSKVRIYEVQATGQSGIKALYBHQGPVLVS 80
QY 63 AF-YDPTHAWSGGLDHLQKMHDLNTDQENLVGTHDAPIRCVCEYCPVNV---VMVTGSMQ 118
Db 81 NWSRDGTKVASGVSDKSAKVFDTQGGQVAAHDDAVRCVRFVEAMGTSPILATGSMQX 140
QY 119 TVKLMDPRTPCNAGTFSQBEKVYTLVSQDRLIVGTAGRRVLWDLRNNGYVQQRRESL 178
Db 141 TLKYWDLRQSTPIATVSLPERVYAMDCVHPLLTAVTAERNICVINLSBPTKIFKLAMSP 200
QY 179 KYQTRCIRAFPNKQGVLSISIEGRVAVELDPSPEVOKKYAFKCHRLKENNIEIYVNV 238
Db 201 KQTRSLACFKDGYAGSVGRCATQNIDEKNASQ--NFSFRCHRNQAGNSADVSVN 258
QY 239 AISFHNHNTFATGSGDGFVNIWDPFNKRLCQFHRYPTSIASLAFSNDGTTLAIASSY 297
Db 259 SIAPHPQYGTFTAGSDGTGTFWDKDSHQLKSPVNVGTTSCSTFNRTGDIPIAYASID 318
QY 298 ---MYEMDDTEHPEDGIFIRQVTDATKPK 324

Db 281 QTRKKIKFAKFNEDSVVKIACSDN--ILCLATS-----DDT 315

RESULT 11

US-09-903-190-157
; Sequence 157, Application US/09903190
; Publication No. US20030162176A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bouquellet, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET.021A
; CURRENT APPLICATION NUMBER: US/09/903,190
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US/09/247,155A
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-04
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 157
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -22..-1
US-09-903-190-157

Query Match 14.6%; Score 258; DB 12; Length 59;

Best Local Similarity 100.0%; Pred. No. 7.4e-20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 MRLKYQHTGAVLDCAFYDPTHAWSGGLDHLKXHDNLNTDOENLVGT 94
|||||
Db 1 MRLKYQHTGAVLDCAFYDPTHAWSGGLDHLKXHDNLNTDOENLVGT 46

RESULT 12

US-09-925-299-1045
; Sequence 1045, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1045
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE

LOCATION: (16)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
LOCATION: (19)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
LOCATION: (23)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
LOCATION: (43)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
LOCATION: (48)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
LOCATION: (49)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
LOCATION: (53)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
LOCATION: (75)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
LOCATION: (78)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
LOCATION: (89)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
LOCATION: (91)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
LOCATION: (99)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-1045

Query Match 11.1%; Score 197.5; DB 9; Length 100;

Best Local Similarity 72.2%; Pred. No. 4.5e-13;
Matches 39; Conservative 2; Mismatches 12; Indels 1; Gaps 1;

QY 44 VPANSMRLKYQHTGAVLDCAFYDPTHAWSGGLDHLKXHDNLNTDOENLVGT 97
|||||
Db 24 VPANSMRLKYQHTGAVLDCXFXGYPXAXWSXGLDHLKXHDNLNTDOENLVGT 76

RESULT 13

US-09-925-299-1045
; Sequence 1045, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1045
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

```

; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (19)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (23)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (43)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (48)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (49)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (53)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (75)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (78)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (89)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (91)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (99)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-299-1045

```

```

Query Match 11.1%; Score 197.5; DB 11; Length 100;
Best Local Similarity 72.2%; Pred. No. 4.5e-13;
Matches 39; Conservative 2; Mismatches 12; Indels 1; Gaps 1;

QY 44 VPANSMRLKYQHTGAVLDCAFYDPTHWAGGGLDQKMDLNTDQENLVGTHDA 97
DB 24 VPANSMRLKYQHTGAVLDCAFYDPTHWAGGGLDQKMDLNTDQENLVGTHDA 76

RESULT 14
US-10-077-111-11
; Sequence 11, Application US/10077111
; Publication No. US20020187492A1
; GENERAL INFORMATION:
; APPLICANT: Toddgerud, C. Gordon
; APPLICANT: Finger, Joshua N.
; APPLICANT: Rillema, Jill
; TITLE OF INVENTION: TBA
; FILE REFERENCE: 3053-4114US2
; CURRENT APPLICATION NUMBER: US/10/077,111
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/294,181
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/269,366
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 11
; LENGTH: 742

```

```

; TYPE: PRT
; ORGANISM: Thermomonospora curvata
; FEATURE:
; OTHER INFORMATION: amino acid sequence encoded by the PKWA gene
US-10-077-111-11

Query Match 11.1%; Score 197.5; DB 14; Length 742;
Best Local Similarity 27.7%; Pred. No. 8.8e-12;
Matches 70; Conservative 48; Mismatches 112; Indels 23; Gaps 10;

QY 15 DGISVSPKPNSTOFLVSSWDTSVRLYDVPANSMRLKYQ-HTGAVLDCAFY-DPTHWAGS 72
DB 502 DMVRAVAFSPD-GALLASGSDATVRLMDVAAERAVFEGHTHVLDIAFSPDGSVMVAS 560
QY 73 GGLDQKMDLNTDQENLV-GTHDAPIRCVCVEYCPVNMVMTGSMWDTVKLMDPRTPCNA 131
DB 561 GSRGTARLNMVATGTGTHAVLKHTDVIYVAFSPDGSVMVASGSRDGTIRLMDVATSKER 620
QY 132 GTFSQP-EKVYTLVSVDG-RLIVGTAGRRVLVMDLRNMGYVQQRRESSLKYQTRCIRAFP 189
DB 621 DVLQAPAEVVVSLAFSPDGSMLVHGSDBSTVHLMDVASGEALH-----TFEGHTDWRAVA 675
QY 190 -NKGYVLSSTEGRAVVEYLDPSPEVQKKYAFKCHRLKENNIQIYFVNAISFHNHT 248
DB 676 FSPDGLASGSDRTIRLMDVAAQEE-----HTTLEGHT-----PVHSVAFHPGTT 724
QY 249 FATGSGDFVNIW 261
DB 725 LASASEDGTIRW 737

RESULT 15
US-10-119-932-1
; Sequence 1, Application US/10119932
; Publication No. US20020160488A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Lal, Preeti
; Guegler, Karl J.
; Corley, Neil C.
; Patterson, Chandra
; TITLE OF INVENTION: HUMAN GTPASE-ASSOCIATED PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Genomics, Inc.
; STREET: 3160 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 3.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/119,932
; FILING DATE: 09-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/590,540
; FILING DATE: 08-Jun-2000
; APPLICATION NUMBER: 09/063,743
; FILING DATE: April 21, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: David G. Streeter
; REGISTRATION NUMBER: 43,188
; REFERENCE/DOCKET NUMBER: PF-0508-1 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:

```

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LUNGN0T01
; CLONE: 059953
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-119-932-1

```

```

Query Match      10.0%; Score 176.5; DB 14; Length 375;
Best Local Similarity 25.8%; Pred. No. 5.5e-10;
Matches 80; Conservative 51; Mismatches 142; Indels 37; Gaps 13;

QY 14 EDGISSVXFSPNTSOFLLVSSMDTSVRLYDVPA---NSMRLKYQHTGAVLDCAP-YDPETH 69
Db 66 EGEVYCCRFHPNGST-LASAGFDRLLLNWVVGDCDNYATLK-GHSGAVMELHYNTDGS 123
QY 70 AWSGGLDHQKMHDLNTDQ--ENLVGTHDAPIRCVEYCEPVNVVMTGSMQDTVKLWDPT 127
Db 124 LFSASTDKTVAWMDSETGERVKRLKGHTSFVNSCYPARRGPQLVCTGSDDDGTGKLWDIRK 183
QY 128 PCNAGTFSQPEKV--YTLVSVDRLIVGTAGRVLVWDLRNMGVYQQRRSSSLKYQTR-- 183
Db 184 KAAIQTFONTYQVLAVTNDTSQIISSGIDNDIKVWDL-----RQNKLTVTWREGH 234
QY 184 ----CIRAPPNKQGYVLSSIEGRVAVEYLDPSPEVQKKYAFKCHRLKENNIEQIYPVNA 239
Db 235 ADSVTGLSLSEGSVLLSNAMDN-TVRYWDVVRPAPKE----RCVKIFQGNVHN-FEKNL 288
QY 240 I--SPHNHNTFATGSGDCFNINDPFNKKRLCQPHRYPTSIASLAFSNDGTTLIAISS- 296
Db 289 LRCSWSPPGSKIAAGSADRFFVYVNDTTSRRILYKLPGHAGSINEVAFHFDPEPIISASSD 348
QY 297 ---VMYEMDD 303
Db 349 KRLYNGEFSE 358

```

Search completed: November 12, 2003, 20:12:03
Job time : 31 secs